Novel algorithmic approaches for the forward and inverse M/EEG problems

Kostiantyn Maksymenko

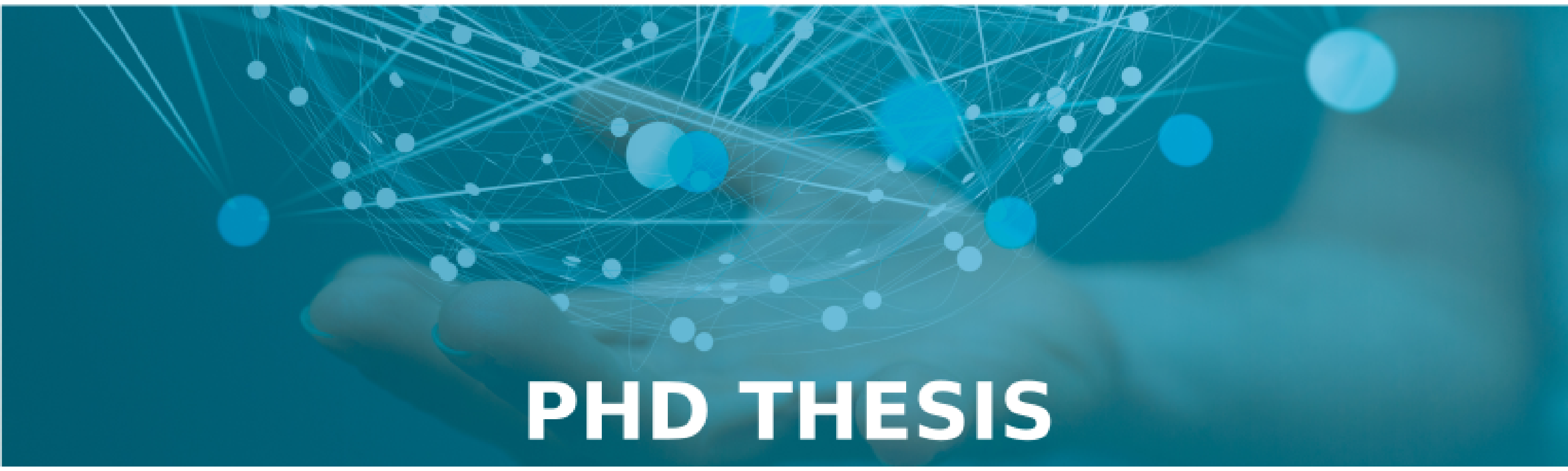
To cite this version:

Kostiantyn Maksymenko. Novel algorithmic approaches for the forward and inverse M/EEG problems. Medical Imaging. Université Côte d'Azur, 2019. English. NNT : . tel-02404166v1

HAL Id: tel-02404166
https://inria.hal.science/tel-02404166v1

Submitted on 11 Dec 2019 (v1), last revised 3 Jul 2020 (v2)

ÉCOLE DOCTORALE
SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION



PHD THESIS

Novel Algorithmic Approaches for the Forward and Inverse M/EEG problems

Kostiantyn MAKSYMENKO
Inria Sophia Antipolis – Méditerranée, Athena Project Team

Submitted in partial fulfillment of the requirements for the degree of Doctor of Science Specialized in Computer Science **of the** Université Côte d'Azur

Advisor : Theodore Papadopoulo
Co-Advisor : Maureen Clerc

Tentative date of defense :
19/12/2019

In front of a Jury composed of:

Alexandre Gramfort, Research Director,
INRIA Saclay - Reviewer
Jens Haueisen, Professor, Technische Universität
Ilmenau, Germany - Reviewer
Théodore Papadopoulo, Research Director,
INRIA Sophia Antipolis - Thesis Advisor
Maureen Clerc, Research Director,
INRIA Sophia Antipolis - Thesis Co-advisor
Juliette Leblond - Research Director,
IINRIA Sophia Antipolis - Examiner
Leandro Beltrachini, Lecturer,
Cardiff University, Wales - Examiner
Christian Bénar, Research Director,
Aix-Marseille Université - Examiner

UNIVERSITÉ
CÔTE D'AZUR

ÉCOLE DOCTORALE
SCIENCES ET
TECHNOLOGIES DE
L'INFORMATION ET DE
LA COMMUNICATION

THÈSE DE DOCTORAT

Nouvelles approches algorithmiques pour les problèmes directs et inverses en M/EEG

Kostiantyn MAKSYMENKO
Inria Sophia Antipolis – Méditerranée, Équipe-Projet Athéna

Présentée en vue de l'obtention du grade de docteur en Informatique
d'Université Côte d'Azur

Dirigée par : Theodore Papadopoulo
Co-encadrée par : Maureen Clerc

Soutenue le :
19/12/2019

Devant le jury, composé de :
Alexandre Gramfort, Directeur de Recherche,
INRIA Saclay - Rapporteur
Jens Haueisen, Professeur, Technische Universität
Ilmenau, Germany - Rapporteur
Théodore Papadopoulo, Directeur de Recherche,
INRIA Sophia Antipolis - Directeur de thèse
Maureen Clerc, Directrice de Recherche,
INRIA Sophia Antipolis - Co-Encadrante de thèse
Juliette Leblond - Directrice de Recherche,
IINRIA Sophia Antipolis - Examinatrice
Leandro Beltrachini, Maître de Conférences,
Cardiff University, Wales - Examinateur
Christian Bénar, Directeur de Recherche,
Aix-Marseille Université - Examinateur

Inria

Résumé

La Magnéto- et l'Electro-encéphalographie (M/EEG) sont deux modalités d'imagerie fonctionnelle non invasives qui mesurent l'activité électromagnétique du cerveau. Ces techniques sont utilisées pour des études cognitives ainsi que pour des applications cliniques, comme l'épilepsie. Après une présentation de quelques notions de base sur ces modalités M/EEG, cette thèse développe deux contributions principales. La première est une méthode d'approximation efficace d'un ensemble de solutions de problèmes directs d'EEG paramétrés par des valeurs de conductivité pour différents tissus. Ce problème direct consiste à calculer comment une activité corticale spécifique serait mesurée par des capteurs EEG. Le principal avantage de notre méthode est qu'elle accélère considérablement le temps de calcul tout en contrôlant l'erreur d'approximation. Les valeurs de conductivités des tissus de la tête varient selon les sujets et il serait intéressant de les estimer à partir des données EEG. Notre méthode est un pas important pour la résolution efficace d'un tel problème d'estimation de conductivités. La deuxième contribution est une nouvelle méthode de reconstruction de sources qui estime des configurations de sources corticales étendues expliquant les mesures M/EEG. La principale originalité de cette méthode réside dans le fait qu'au lieu de fournir une reconstruction unique, comme le font la majorité des méthodes de l'état de l'art, elle propose plusieurs solutions candidates valables. Nous avons validé nos deux contributions sur des données M/EEG simulées et réelles.

Mots-clés EEG, MEG, problème inverse, problème direct, reconstruction de source, estimation de conductivité.

Abstract

Magneto- and Electro-encephalography (M/EEG) are two non-invasive functional imaging modalities which measure the electromagnetic activity of the brain. These tools are used in cognitive studies as well as in clinical applications as, for example, epilepsy. Besides the presentation of some background material about the M/EEG modalities, this thesis describes two main contributions. The first one is a method for a fast approximation of a set of EEG forward problem solutions, parametrized by tissue conductivity values. This forward problem consists in computing how a specific cortical activity would be measured by EEG sensors. The main advantage of our method is that it significantly accelerates the computation time, while controlling the approximation error. Head tissue conductivity values vary across subjects and it might be interesting to estimate them from the EEG data. Our method is an important step towards an efficient solution of such a head tissues conductivity estimation problem. The second contribution is a novel source reconstruction method, which estimates extended cortical sources explaining the M/EEG measurements. The main originality of the method is that instead of providing a unique reconstruction, as the majority of the state-of-the-art methods do, it proposes several equally valid candidates. We validated both our contributions on simulated and real M/EEG data.

Key words: EEG, MEG, inverse problem, forward problem, source reconstruction, conductivity estimation.

Acknowledgements

This work was supported by ANR grant VIBRATIONS (ANR-13-PRTS-0011) and by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (ERC Advanced Grant agreement No 694665 : CoBCoM - Computational Brain Connectivity Mapping).

Contents

General introduction

Context

The last frontier of the biological sciences – the ultimate challenge – is to understand the biological basis of consciousness and the brain processes by which we feel, act, learn, and remember [Kandel et al., 2012].

One of the fundamental objectives of neuroscience is to determine the link between a particular mental process (perception, movement, thought, etc.) and functional activity of specific brain regions, as well as their anatomy and physiology. Different medical imaging techniques are used to study this relationship..

Neuroimaging modalities

Structural neurimaging deals with the structure of the nervous system. The study of brain activity with medical imaging methods is called functional neuroimaging. Her we give a brief list of neuroimaging techniques that are used nowadays. Magnetic Resonance Imaging (MRI) provides anatomical images (description of the brain in terms of tissues and macroscopic model of brain shape). Functional MRI (fMRI) measures brain activity by detecting changes associated with blood oxygenation [Ogawa et al., 1990]. Diffusion MRI (dMRI) provides a unique picture on brain anatomical connectivity by measuring the anisotropy in white matter tracts [Poupon et al., 2001]. Computed tomography (CT) scanning of the head uses a series of x-rays of the head taken from many different directions; the resulting data is transformed into a series of cross sections of the brain using a computer program. Positon Emission Tomography (PET, SPECT) is used to observe metabolic processes in the brain [Nasrallah and Dubroff, 2013]. Magneto- and Electro-encephalography (MEG/EEG) perform measurement of the electromagnetic activity of the brain on the scalp (EEG) or outside the head (MEG). Stereo-electroencephalography (SEEG) is a method for recording electroencephalographic signals via deop electrodes (electrodes surgically implanted into the brain tissue) [Talairach, 1974, Youngerman et al., 2019]. Near-infrared spectroscopy (NIRS), a spectroscopic method that uses the near-infrared region of the electromagnetic spectrum (from 780 nm to 2500 nm).

All these methods have different spatial and temporal resolutions, as can be seen in Fig. 1.

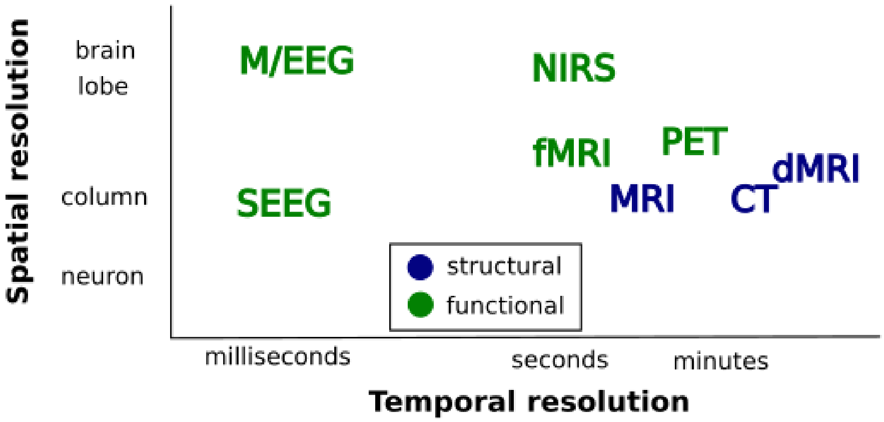


Figure 1: The different modalities for brain imaging. MRI, fMRI, dMRI and CT have the unique property to yield high-resolution spatial information of the whole brain. M/EEG and SEEG provide much better temporal resolution then NIRS or fMRI. The spatial resolution of M/EEG is, however, low. Even though SEEG provides much better spatial resolution, it is an invasive method and it is used only in particular circumstances (e.g. in patients with epilepsy not responding to medical treatment).

Main contributions of the thesis

In this work, we will mostly focus on the M/EEG modalities. They uniquely provide a good temporal resolution of several milliseconds. This allows to capture rapid brain activity, which is impossible with other functional modalities such as fMRI or NIRS. They are also non-invasive methods (unlike SEEG) as the signal is recorded by sensors located on or outside the head. This results, however, in a low spatial resolution. In fact, each sensor measures the signal originating from a cortical area of 10cm^2 or more.

The usual framework with M/EEG starts with designing, setting up an experiment and following data acquisition. Even though it contains a functional information about brain activity, because of the weak spatial resolution, it is not possible to directly localize the origins of the M/EEG signals on the cortex. Additional techniques are required to solve this problem. Computing how a specific cortical activity is measured by M/EEG sensors is called the ***forward problem***. Reconstructing the cortical activity from the M/EEG measurements is called the ***inverse problem***. Fig. 2 shows the general pipeline of working with M/EEG modalities.

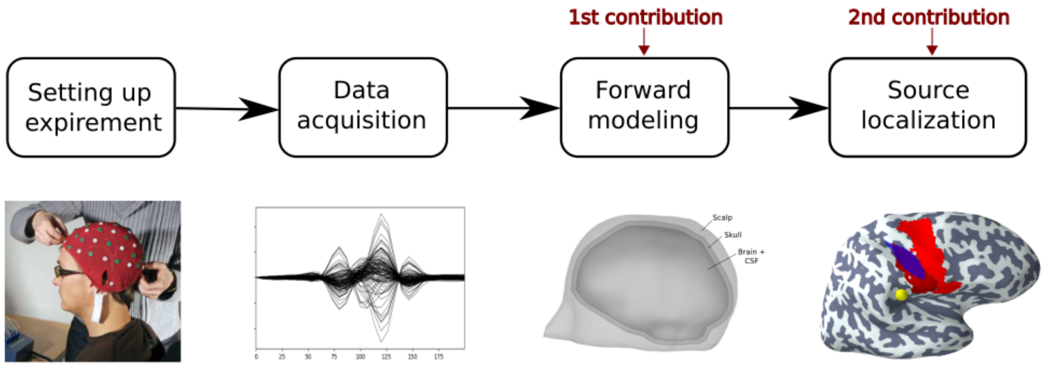


Figure 2: General pipeline of working with M/EEG modalities. M/EEG signal measured during the experiment cannot be directly mapped onto the cortex. Additional techniques are required to solve this problem. First, ***forward model*** should be computed, which is then used to solve the source localization (***inverse***) problem.

In this thesis we present two main contributions. On of them is mostly related to the EEG forward problem. Our second contribution is a novel method for solving the M/EEG inverse problem. Both of them, however, have a common general objective – improving the accuracy of the M/EEG source localization.

First contribution. Fast approximation of EEG forward problem and application to tissue conductivity estimation.

To compute the M/EEG forward problem, a head volume conductor model (i.e. the distinction of the different conductive compartments in the head) is required. It consists of a volume head geometry and the information about different head tissue conductivities. The geometrical 3d model of the head can include scalp, skull, cortex, white matter fibers, cerebrospinal fluid (CSF), etc. This modeling is mostly based on MRI scans of subjects (or CT for hard tissues, like bone).

It is impractical or impossible to directly measure head tissue conductivities in vivo. That is why, in practice, default conductivity values are often used. These values, however, vary a lot across subjects. For MEG, exact conductivity values are not so important. For EEG, however, using wrong conductivity values may lead to the significant inaccuracy of the forward problem solution. A possible solution to this problem is to estimate unknown conductivities directly from the EEG data. To do so, the EEG forward problem solution must be computed for a large number of conductivity configurations, which may be computationally expensive and impractical.

To overcome this computational problem, we will propose a fast method to ***approximate*** the EEG forward solution for any conductivity configuration. We will show that these approximations can be used for tissue conductivities estimation.

This will in turn lead to a more accurate forward operator (solution of the forward problem).

Second contribution. Providing a family of solutions to the M/EEG source reconstruction problem through data-driven cortical clustering. The M/EEG inverse problem is, in general, ill-posed. In particular, it means that measurements and forward operator alone are not sufficient to find a unique solution. In other words, given a set of measurements from M/EEG sensors, an infinite number of cortical activity configurations can explain it. Thus, additional a priori hypotheses about the cortical activity are needed to constrain the solution. Different source reconstruction methods were proposed to overcome this problem.

The most of the state-of-art inverse methods aim to reconstruct a single cortical source configuration which explains the data. However, because of the ill-posedness of the problem, it is very likely that other spatially distinct source configurations could explain the data equivalently well as the identified solution. That is why we think, that the methods which are forced to find a unique solution, without giving alternative ones, can be disadvantageous, because they may find a false solution.

To overcome this issue, in this thesis, we propose a new approach for M/EEG source reconstruction. It is based on the idea that it is better to have several possible candidates to a solution, with the true one among them, than to have a single, but possibly false solution. That is why our method provides several candidate solutions. These candidates are different in term of their extent and/or positions on the cortex, but fit the data with similar accuracy.

In spite of some limitations of our approach, we believe that the innovative features that our method provides will make it useful to the M/EEG community.

Overview of the thesis

Chapter 1. Neurophysiological Bases of EEG and MEG. This chapter introduces the neurophysiological origins of EEG and MEG signals. Starting with the description of the electro-magnetic activity of a single neuron, we describe how it can be scaled up to the activity of cortical regions, and, finally, how it cab by measured be M/EEG sensors. We analyse the main similarities and dissimilarities of EEG and MEG. This introductive chapter gives the nneurophysiological grounding for the mathematical formulation of the M/EEG forward and inverse problems, presented in the following chapters.

Chapter 2. Forward modeling of EEG and MEG signals. This chapter provides a mathematical formulation of the M/EEG forward problem. Starting by introducing fundamental equations which describe the electro-magnetic fields generated by cortical activity, we then describe the head and source modeling. This leads to the definition of the forward operator – lead field matrix – which maps the cortical activity to the M/EEG sensor measurements. The content of this chapter is crucial for understanding the following chapters, especially the Chapter 4, which is very related to the M/EEG forward problem.

Chapter 3. M/EEG inverse problem. This chapter gives an overview of several classes of state-of-art M/EEG source localization methods. Starting with the general formulation of well- and ill-posed problems, we introduce basic principles of dipole fit and scanning methods, as well as variational and Bayesian approaches to solve the M/EEG inverse problem. What is important is the MUSIC algorithm, because our approach, that we introduce in Chapter 5, is based on this method.

Chapter 4. Fast Approximation of EEG Forward Problem and Application to Tissue Conductivity Estimation. This chapter introduces our first contribution – a method for a fast approximation of EEG forward problem solution. We provide a theoretical description of the algorithm and test its performance on simulated and real EEG data. This chapter is based on a published article [Maksymenko et al., 2019].

Chapter 5. Providing a family of solutions to the M/EEG source reconstruction problem through data-driven cortical clustering. This chapter introduces our second contribution – a new M/EEG source reconstruction method that we named CLUstering Based MUSIC (CLUB-MUSIC). This method is based on a M/EEG data-driven clustering of cortical sources. It is an extension of the state-of-art MUSIC algorithm, adapted for sources represented by cortical regions. Another feature of our method is that it provides not a single solution, but several candidate solutions. These candidates represent spatially distinct cortical regions and fit the data with similar accuracy.

Chapter 6. Results of CLUB-MUSIC on simulated and real MEG data. This chapter evaluates our CLUB-MUSIC method on simulated and real MEG data. Starting with a simulation of a single region, we then test our method on a multiple region simulation and, finally, on real MEG data corresponding to the auditory evoked fields. The results show that, indeed, several spatially distinct source configurations can explain the data with similar accuracy. Our method is able to provide several candidates for the solution. It is based on the principle that it is better to obtain a few false positives regions, than a single false negative one.

General conclusion. This chapter concludes the main contribution of this thesis. We also discuss the current limitation of proposed methods and the main perspectives for the future work.

Author's publications and other contributions

Peer-reviewed publications:

- **Kostiantyn Maksymenko**, Maureen Clerc, Théodore Papadopoulo. Fast Approximation of EEG Forward Problem and Application to Tissue Conductivity Estimation. IEEE Transactions on Medical Imaging, Institute of Electrical and Electronics Engineers, In press, (10.1109/TMI.2019.2936921).
- **Kostiantyn Maksymenko**, Maureen Clerc, Théodore Papadopoulo. Data-driven cortical clustering to provide a family of plausible solutions to M/EEG inverse problem. iTWIST, Nov 2018, Marseille, France. (hal-01946876).
- **Kostiantyn Maksymenko**, Giusiano B., Roehri N., Bénar CG, Badier JM, Strategies for statistical thresholding of source localization maps in magnetoencephalography and estimating source extent, Journal of Neuroscience Methods, Volume 290, pp. 95-104, Oct 2017, https://doi.org/10.1016/j.jneumeth.2017.07.015

Participation in the conferences:

- **Kostiantyn Maksymenko**, Maureen Clerc, Théodore Papadopoulo. Data-driven cortical clustering to provide a family of plausible solutions to the M/EEG inverse problem. BIOMAG 2018, Aug 2018, Philadelphia, United States. (hal-01874281)
- **Kostiantyn Maksymenko**, Maureen Clerc, Théodore Papadopoulo. A fast EEG forward problem approximation method and its application to tissue conductivity estimation. BACI 2017, Aug 2017, Bern, Switzerland.

Collaboration with other authors:

- **Kostiantyn Maksymenko**, Théodore Papadopoulo, Samuel Deslauriers-Gauthier. White matter fiber bundles as a source model in the MEG inverse problem. COBCOM 2017 - Computational Brain Connectivity Mapping, Winter School Workshop, Nov 2017, Juan-les-Pins, France. 2017. (hal-01713075)
- Isa Costantini, Patryk Filipiak, **Kostiantyn Maksymenko**, Samuel Deslauriers-Gauthier, Rachid Deriche. fMRI Deconvolution via Temporal Regularization using a LASSO model and the LARS algorithm. EMBC'18 - 40th International Engineering in Medicine and Biology Conference, Jul 2018, Honolulu, United States. (hal-01855467)

Other experiences and contributions:

- **iTWIST'18**, doctoral school and workshop on interactions between low-complexity data models and sensing techniques, `https://sites.google.com/view/itwist18/workshop`.
- **Summer School**, inverse problems and imaging, `http://www.math.uni-bremen.de/zetem/cms/detail.php?template=ipschool2017_parse_title&person=ip-school2017`.
- **MNE-Python coding sprint**, one of the initiators and developers of a simulation module of MNE-Python package, `https://mne-tools.github.io/sprint2019/`.

Bibliography

Eric R. Kandel, James H. Schwartz, Thomas M. Jessel, Steven A. Siegelbaum, and A. J. Hudspeth. *Principles of neural science*. 5 edition, 2012. ISBN 9780071390118.

K. Maksymenko, M. Clerc, and T. Papadopoulo. Fast approximation of eeg forward problem and application to tissue conductivity estimation. *IEEE Transactions on Medical Imaging*, 2019. ISSN 0278-0062. doi: 10.1109/TMI.2019.2936921.

Ilya Nasrallah and Jacob Dubroff. An overview of pet neuroimaging. *Seminars in Nuclear Medicine*, 43(6):449 – 461, 2013. ISSN 0001-2998. doi: https://doi.org/10.1053/j.semnuclmed.2013.06.003. URL `http://www.sciencedirect.com/science/article/pii/S0001299813000561`. Non-Oncology PET (Part II).

S Ogawa, T M Lee, A R Kay, and D W Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, 87(24):9868–9872, 1990. ISSN 0027-8424. doi: 10.1073/pnas.87.24.9868. URL `https://www.pnas.org/content/87/24/9868`.

C. Poupon, J.-F. Mangin, C.A. Clark, V. Frouin, J. Régis, D. Le Bihan, and I. Bloch. Towards inference of human brain connectivity from mr diffusion tensor data. *Medical Image Analysis*, 5(1):1 – 15, 2001. ISSN 1361-8415. doi: https://doi.org/10.1016/S1361-8415(00)00030-X. URL `http://www.sciencedirect.com/science/article/pii/S136184150000030X`.

J. Talairach. *Approche nouvelle de la neurochirurgie de l'épilepsie: Méthodologie stéréotaxique et résultats thérapeutiques*. Neuro-Chirurgie. Masson, 1974. URL `https://books.google.fr/books?id=1Q3-PgAACAAJ`.

Brett E Youngerman, Farhan A Khan, and Guy M McKhann. Stereoelectroencephalography in epilepsy, cognitive neurophysiology, and psychiatric disease: safety, efficacy, and place in therapy. *Neuropsychiatric Disease and Treatment*, Volume 15:1701–1716, jun 2019. ISSN 1178-2021. doi: 10.2147/NDT.S177804.

Chapter 1

Neurophysiological Bases of EEG and MEG.

Contents

1.1. Overview

This chapter is an introduction of the magnetoencephalography (MEG) and electroencephalography (EEG) modalities. We cover the basic aspects of cellular composition, morphology and function of the human brain. Then, we describe the structure and function of neurons. We explain the basic electrico-chemical processes which allow neurons to generate and spread signals. Finally, we describe the origins of the EEG and MEG measurements. Further readings on these subjects can be found in the following books: Electroencephalography [Niedermeyer and Da Silva, 2004] ; Principles of Neural Science [Kandel et al., 2012] ; MEG-EEG Primer [Hari et al., 2017].

1.2. Introduction to EEG and MEG

Magnetoencephalography (MEG) and electroencephalography (EEG) (M/EEG) are two complementary techniques that measure, respectively, the magnetic field (outside the head) and the distribution of electric potentials on the scalp produced by electrical activity in brain neural cell assemblies [Baillet et al., 2001]. Among the available functional imaging techniques, MEG and EEG uniquely have temporal resolutions of several milliseconds. Thus, with M/EEG it is possible to follow the rapid changes in cortical activity that reflect ongoing signal processing in the brain; the electrical events of single neurons typically last from one to several tens of milliseconds [Hämäläinen et al., 1993].

1.2.1. EEG

EEG consists of measurements of a set of electric potential differences between pairs of scalp electrodes. Sensors may be either directly glued to the skin (for prolonged clinical observation) at selected locations directly above cortical regions of interest or fitted in an elastic cap for rapid attachment with near uniform coverage of the entire scalp (Fig. 1.1). Research protocols can use up to 256 electrodes [Baillet et al., 2001]. Typical electric potentials on the scalp are in the range 1-100 mV. To avoid external electrical interference, EEG is preferably measured inside a Faraday cage that dampens power-line artifacts and other electrical noise, although recording of sufficient quality can also be performed in regular or operating rooms and even in mobile real-life settings using mobile EEG devices [Hari et al., 2017].

The scalp electroencephalogram, recorded by a single electrode, is a spatially smoothed version of the local field potential (LFP), integrated over an cortical area of 10 cm^2 or more. It has little discernible relationship with the firing patterns of the contributing individual neurons, and this is largely due to the distorting and attenuating effects of the soft and hard tissues between the current source and the recording electrode [Buzsáki et al., 2012].

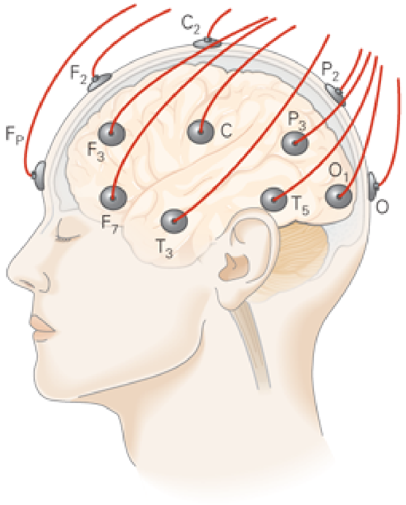


Figure 1.1: Schematic representation of EEG electrodes placed on the scalp.. Adapted from [Kandel et al., 2012].

1.2.2. MEG

Magnetic signals from brain sources are extremely weak compared with ambient magnetic field and its variations. Thus neuromagnetic signals are typically 50-500 fT, one part in 10^9 or 10^8 of the earth's geomagnetic field. The detector that offers sufficient sensitivity for the measurement of these tiny fields is the superconducting quantum interference device (SQUID). The sensitivity of a SQUID measuring system to external magnetic noise is greatly reduced by the proper design of the flux transformer (Fig.1.2). Systems of coils such as gradiometers are insensitive to spatially uniform changes in the background field, but respond to inhomogeneous changes. MEG measurements are also usually performed in a magnetically shielded room [Hämäläinen et al., 1993]. However, SQUIDs are usually located several centimeters away from the scalp because of the cryogenic needs, which decreases the signal to noise ratio (SNR) of the data.

Recent advances in the field of quantum technology have led to the development of small, optically-pumped magnetometers (OPMs). These devices measure the transmission of laser light through a vapour of spin-polarised atoms, which provides a highly sensitive measure of the local magnetic field. OPMs have theoretical sensitivity comparable to that of the SQUIDs used in current MEG systems, but operate without cryogenic cooling (the vapour cell is sometimes heated, but the external surface remains close to room temperature). This means it is now

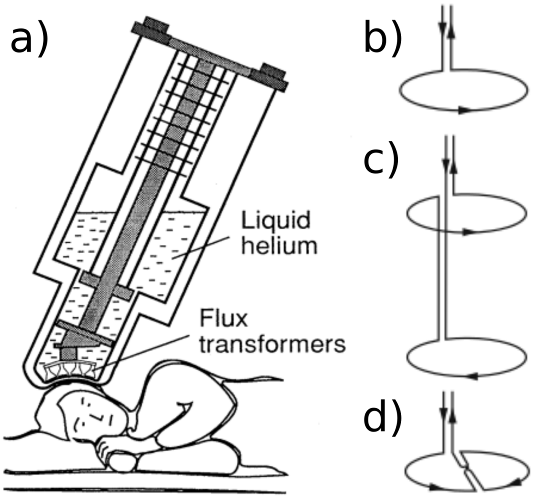


Figure 1.2: Detection of cerebral magnetic fields. a) Schematic representation of the MEG scanner. The bottom of the helium dewar, with the flux-transformer pickup coils near subject's head, is brought as close to the head as possible. Three superconducting flux transformers employed in brain research: b) magnetometer c) axial gradiometer, d) planar gradiometer. Adapted from [Hämäläinen et al., 1993].

possible to consider a MEG system with sensitive detector volumes just 6 mm from the scalp surface. This technology is currently being commercialized and empirical performance is increasing, which shows the potential of OPMs to transform the MEG measurements. This promises not only better SNR, but also access to traditionally challenging subject groups such as infants [Boto et al., 2017].

1.3. Structure of neurons

The cell body or soma is the metabolic center of the neuron. It contains the nucleus, which contains the genes of the cell, and the endoplasmic reticulum, an extension of the nucleus where the cell's proteins are synthesized. The cell body usually has to two kinds of extensions: several short dendrites and one long, tubular axon. Dendrites branch out in tree-like fashion and are the main apparatus for receiving incoming signals from other nerve cells. The axon typically extends some distance from the cell body and transmits signals to other neurons. An axon can convey electrical signals over distances ranging from 0.1 mm to 2 m. These electrical signals, called action potentials, are initiated at a specialized trigger region near the origin of the axon called the initial segment from which they propagate down the axon without failure or distortion at speeds of 1 to 100 m/s. The amplitude of an action potential traveling down the axon remains constant at 100 mV because the action potential is an all-or-none impulse that is regenerated at regular intervals along the axon. To increase the speed by which action potentials are conducted, large axons are wrapped in an insulating sheath of a lipid substance, myelin. The sheath is interrupted at regular intervals by the nodes of Ranvier, uninsulated spots on the axon where the action potential is regenerated. Near its end, the axon divides into fine branches that contact other neurons at specialized zones of communication known as synapses. The nerve cell transmitting a signal is called the presynaptic cell; the cell receiving the signal is the postsynaptic cell. The presynaptic cell transmits signals from specialized enlarged regions of its axon's branches, called presynaptic terminals or nerve terminals. The presynaptic and postsynaptic cells are separated by a very narrow space, the synaptic cleft. Most presynaptic terminals end on the postsynaptic neuron's dendrites; but the terminals may also terminate on the cell body or, less often, at the beginning or end of the axon of the postsynaptic cell. Fig. 1.3 show the schematic representation of the structure of a neuron.

1.4. Signal producing by neuron cells

Regardless of cell size and shape, transmitter biochemistry, or behavioral function, almost all neurons can be described by a model neuron that has four functional components that generate the four types of signals: a receptive component, a summing or integrative component, a long-range signaling component, and a secretory component.

The different types of signals generated in a neuron are determined in part by the electrical properties of the cell membrane. All cells, including neurons, maintain a certain difference in the electrical potential on either side of the plasma membrane when the cell is at rest. This is called the resting membrane potential. In a typical resting neuron, the voltage of the inside of the cell is about 65 mV below the voltage outside the cell. Because the voltage outside the membrane is defined as zero, we say the resting membrane potential is –65 mV. The resting membrane potential results from the selective permeability of the membrane and unequal distribution of electrically charged ions, in particular the positively charged Na^+ , and K^+ ions.

A cell is said to be excitable when its membrane potential can be quickly and significantly altered. This change serves as a signaling mechanism. In some neurons reducing the membrane potential by 10 mV (from –65 to –55 mV) makes the membrane much more permeable to Na^+ than to K^+. The resultant influx of positively charged Na^+ neutralizes the negative charge inside the cell and causes a brief and explosive change in membrane potential to +40 mV. This ***action potential*** is conducted down the cell's axon to the axon's terminal, where it

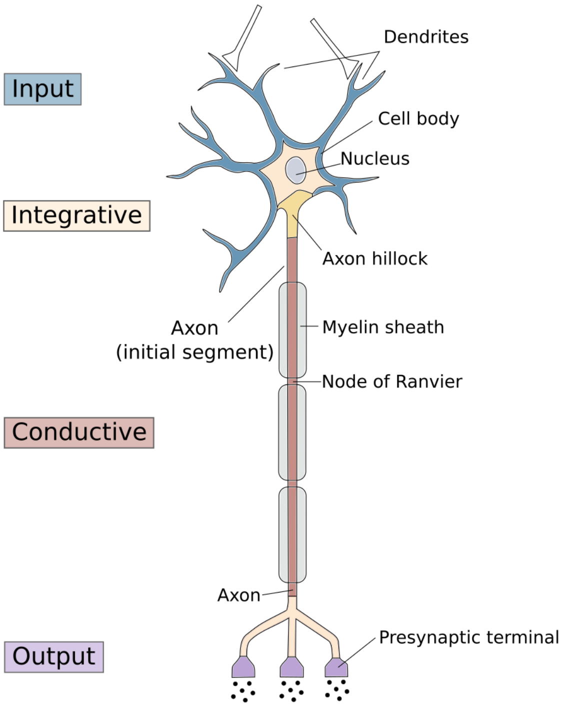


Figure 1.3: The structure of a neuron. Most neurons in the vertebrate nervous system have several main features in common. The cell body contains the nucleus, the storehouse of genetic information, and has two types of cell extensions: axons and dendrites. Axons are the transmitting element of neurons. Many axons are insulated by a sheath of fatty myelin that is regularly interrupted at gaps called nodes of Ranvier. Most neurons also have four functional regions, in which different types of signals are generated. The input, integrative, and conductive signals are all electrical and integral to the cell, whereas the output signal is a chemical substance ejected by the cell into the synaptic cleft. Adapted from [Kandel et al., 2012].

initiates an elaborate chemical communication with other neurons. The action potential is actively propagated along the axon so that its amplitude does not diminish by the time it reaches the axon terminal. An action potential typically lasts approximately 1 ms, after which the membrane returns to its resting state, with its normal separation of charges and higher permeability to K^+ than to Na^+.

The change in membrane potential that generates long-range and local signals can be either a decrease or an increase from the resting potential. The resting membrane potential therefore provides the baseline on which all signaling occurs. A reduction in the potential difference outside and inside the cell is called depolarization. Because depolarization enhances a cell's ability to generate an action potential, it is an excitatory process. In contrast, an increase in in the potential difference is called hyperpolarization. Hyperpolarization makes a cell less likely to generate an action potential and is therefore an inhibitory process.

Input component

A presynaptic terminal releases a chemical substance, or neurotransmitter, into the synaptic cleft. After diffusing across the cleft, the transmitter binds to receptor proteins in the membrane of the postsynaptic neuron, thereby directly or indirectly opening ion channels. The ensuing flow of ions alters the membrane potential of the cell, a change called the synaptic potential (Fig. 1.4). The synaptic potential is graded, i.e. its amplitude depends on how much transmitter is released. In the same cell, the synaptic potential can be either depolarizing or hyperpolarizing depending on the type of receptor molecule that is activated and on the type of transmitter released by the presynaptic cell. Synaptic potentials spread passively and thus are local changes in potential unless the signal reaches beyond the axon's initial segment and thus can give rise to an action potential.

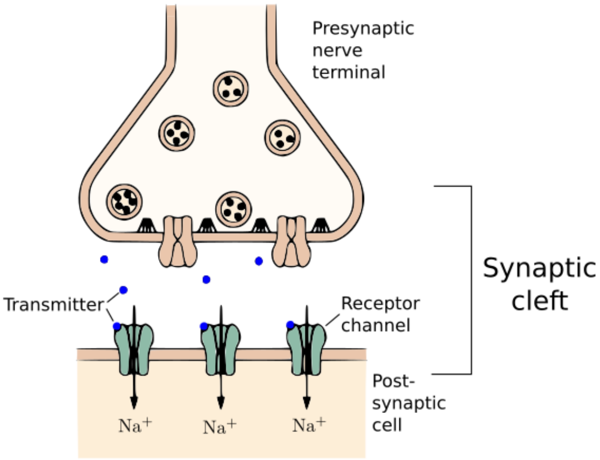


Figure 1.4: Released neurotransmitter molecules diffuse across the synaptic cleft and bind specific receptors on the postsynaptic membrane. These receptors cause ion channels to open (or close), thereby changing the membrane potential of the postsynaptic cell. In this figure, the excitatory synapse activity is illustrated, which results of membrane depolarization. Adapted from [Kandel et al., 2012].

Integrative component

Action potentials are generated by a sudden influx of Na^+ ions through channels in the cell membrane that open and close in response to changes in membrane potential. When an input signal (synaptic potential) depolarizes a local region of membrane, the local change in membrane potential opens local Na^+ channels that allow Na^+ to flow down its concentration gradient, from outside the cell where the Na^+ concentration is high towards inside where it is low.

Because the initial segment of the axon has the highest density of voltage-sensitive Na^+ channels and therefore the lowest threshold for generating an action potential, an input signal spreading passively along the cell membrane is more likely to give rise to an action potential at the initial segment than at other sites in the cell. This part of the axon is therefore known as the trigger zone. It is here that the activity of all synaptic potentials is summed and where, if the sum of the input signals reaches threshold, the neuron generates an action potential.

Conductive component

The action potential is all-or-none: stimuli below the threshold do not produce a signal, but stimuli above the threshold all produce signals with the same amplitude. Whatever the stimuli vary in intensity or duration, the amplitude and duration of each action potential are essentially the same. In addition, unlike receptor and synaptic potentials, which spread passively and decrease in amplitude, the action potential does not decay as it travels along the axon to its target, a distance which can be as great as 2 m, because it is periodically regenerated. This conducting signal can travel at rates as fast as 100 m/s.

The output component

When an action potential reaches a neuron's terminal it stimulates the release of chemical substances from the cell. These substances, called neurotransmitters, can be small organic molecules, such as L -glutamate and acetylcholine, or peptides like substance P or LHRH (luteinizing hormone releasing hormone).

Once released, the neurotransmitter is the neuron's output signal. Like the input signal, it is graded. The amount of transmitter released is determined by the number and frequency of action potentials that reach the presynaptic terminals. After release the transmitter diffuses across the synaptic cleft and binds to receptors on the postsynaptic neuron. This binding causes the postsynaptic cell to generate a synaptic potential. Whether the synaptic potential has an excitatory or inhibitory effect depends on the type of receptor in the postsynaptic cell, not on the particular chemical neurotransmitter. The same transmitter substance can have different effects at different receptors.

Fig. 1.5 shows the signals generated and transmited be different parts of a neuron.

1.5. The origin of EEG signals

In the previous section we briefly introduced the principal signals generated by a neuron. In this section, we will describe more in details the extracellular field

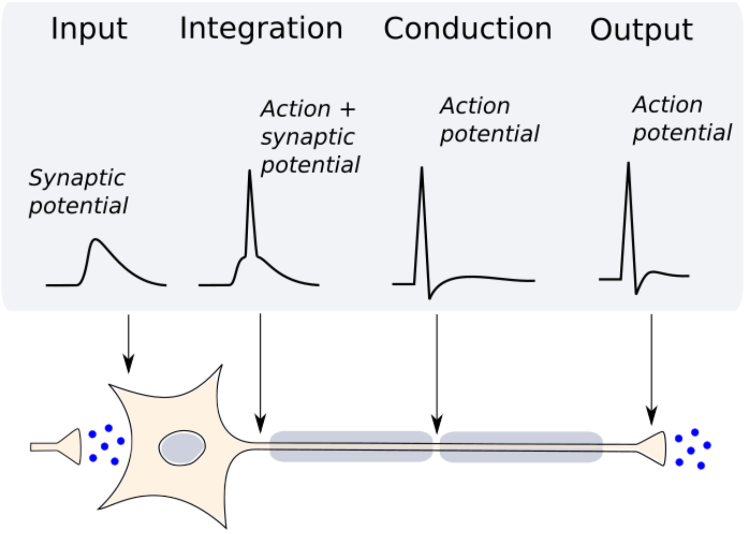


Figure 1.5: The sequence of signals produced by neuron activity. A presynaptic neuron releases excitatory neurotransmitters which initiates a synaptic potential that spreads passively to the trigger zone of the neuron's axon, where, if passing threshold, it initiates an action potential that propagates actively to the terminal of the axon. The action potential releases a neurotransmitter into the synaptic cleft. Let us remark that action potentials are stronger but faster than synaptic potentials. Adapted from [Kandel et al., 2012].

potentials that can be recorded in the neighborhood of their cellular sources, commonly called ***local field potentials*** (LFPs).

Any excitable membrane — whether it is a dendrite, soma, axon or axon terminal — and any type of transmembrane current contributes to the extracellular field. The field is the ***superposition*** of all ionic processes. All currents in the brain superimpose at any given point in space to yield LFP value at that location. However, in the context of M/EEG, synaptic activity is often the most important source of extracellular current flow. The idea that synaptic currents contribute to the LFP stems from the recognition that extracellular currents from many individual compartments must overlap in time to induce a measurable signal, and such overlap is most easily achieved for relatively slow events, such as synaptic currents [Buzsáki et al., 2012] (see Fig. 1.5). Let us notice that even though fast action potentials generate the strongest currents across the neuronal membrane and can be detected as 'unit' or 'spike' activity in extracellular medium, until recently they were thought not to contribute substantially to the traditionally considered LFP band (<100 Hz) or to the scalp-recorded EEG, because the strongest fields they generate are of short duration (<2 ms) and nearby neurons rarely fire synchronously in such short time windows under physiological conditions. However, synchronous action potentials from many neurons can contribute substantially to

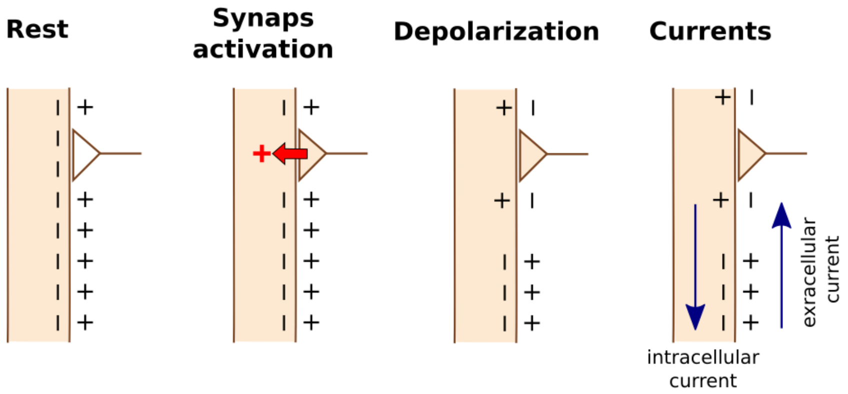


Figure 1.6: Basic mechanism underlying the generation of intra- and extracellular currents after synaptic activity. The description is based on the assumption that the dendrite is locally depolarized by activation of an excitatory synapse. 1) At rest, the membrane potentials of a neuron is about −60 mV. 2) Activation of an excitatory synapse causes the flow of positive ions into the cell. 3) The result is a local change of membrane potential (depolarization). 4) The potential difference between the near synaptic area and the rest of the cell causes intracellular and extracellular currents. Adapted from [Niedermeyer and Da Silva, 2004].

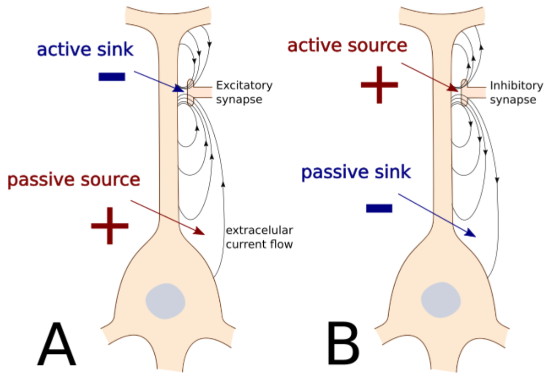


Figure 1.7: Current flow patterns around an idealized neuron owing to synaptic activation. **A)** Current flow caused by the activation of an excitatory synapse at the level of a dendrite, results in the depolarization of the membrane and flow of a net positive current towards inside the cell. This current causes an active sink at the synapse site outside the cell. At the soma level there exists a distributed passive source resulting in an extracellular potential of positive polarity. **B)** Current flow caused by the activation of an inhibitory synapse, results in hyperpolarization of the membrane and flow of a net positive current outwards from the cell. This results in a active source at the level of synapse and passive sink at soma level. Adapted from [Niedermeyer and Da Silva, 2004].

high- frequency components of the LFP [Buzsáki et al., 2012].

In this section, we will discuss the LFP produced by post-synaptic currents as the most important contributor to EEG measurements. As mentioned in a section 1.4, the release of neurotransmitters by presynaptic neuron triggers an ionic current in or out the neuron. It results in local depolarization or hyperpolarization of the postsynaptic membrane. This creates a potential gradient and currents along membrane inside and outside the cell. Fig. 1.6 gives a schematic representation of this process.

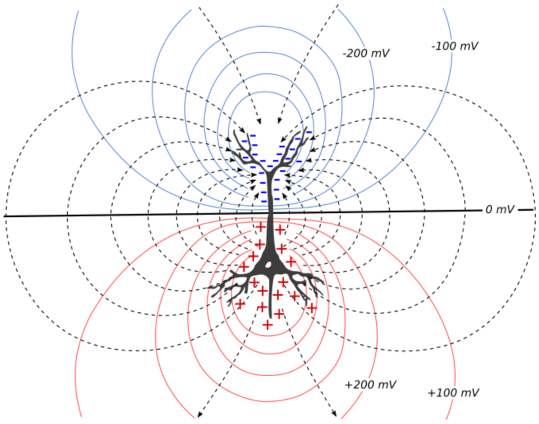


Figure 1.8: Electrical field conforming to that of a dipole created by the synaptic excitation of a single cortical pyramidal neuron. It is assumed that the neuron has been excited by a set of excitatory synapses located on the apical dendrites. Because of the depolarization of the dendritic membrane, this segment of the neuron becomes extracellulary electronegative with respect to the soma and basal dendrites, which become electropositive. This causes current to flow through the extracellular medium as indicated by the black dashed lines with arrows. The potential distribution within the volume conductor corresponding to this current flow is portrayed by the isopotential surfaces represented as colored solid lines. Each isopotential surface represents the geometric locus of all points that are at the same potential. Adapted from [Gloor, 1985].

In the case of excitatory synaptic activity, the positive ionic current is directed to the intracellular medium. It creates an ***active sink*** at the extracellular synaptic level. It results in the presence of a ***passive source*** along the cell and at a distance from the synaptic level (Fig. 1.7a). In the case of an inhibitory postsynaptic potentials, the positive ionic current is directed towards the extracellular medium. It results in an ***active source*** at the synaptic level and a ***passive sink*** along the cell soma and at a distance from the synaptic level (Fig. 1.7b) [Niedermeyer and Da Silva, 2004].

The potential difference causes a current to flow through the volume conductor between the source and the sink. Some of this current takes the shortest route, but current also flows through more distant parts of the volume conductor. Even though current density rapidly drops off with increasing distance from its source, some current, at least theoretically, flows even through the most distant part of the volume conductor [Gloor, 1985].

Based on this sink-source origin of LFP, the electrical field around a single neuron is often modeled as a ***dipole*** [Niedermeyer and Da Silva, 2004, Gloor, 1985, Lindén et al., 2010]. Fig. 1.8 shows the electric field generated by such an idealized "dipole neuron".

EEG electrodes measure these potentials on the scalp surface. Let us remark that the origin or polarity of cortical synaptic events cannot be determined from surface EEG recordings alone. Excitatory postsynaptic potentials (EPSPs) in superficial layers and inhibitory postsynaptic potentials (IPSPs) in deeper layers both appear as upward (negative) potentials, whereas EPSPs in deeper layers and IPSPs in superficial layers have downward (positive) potentials (Fig.1.9).

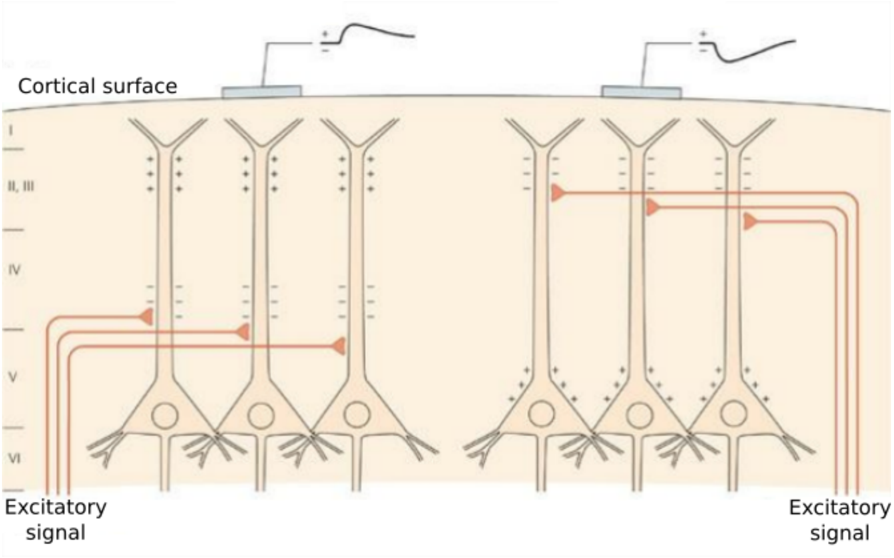


Figure 1.9: Surface electroencephalogram (EEG) recordings do not unambiguously indicate the polarity of synaptic events. The polarity of the surface EEG depends on the location of the synaptic activity within the cortex. An excitatory signal in cortical layer V causes a upward voltage deflection at the surface EEG electrode because the electrode is nearer the current source. In contrast, an excitatory signal in cortical layer II causes an downward deflection because the electrode is nearer the sink. Taken from [Kandel et al., 2012].

1.6. Origin of MEG signals

In the previous section we described the extracellular potentials caused by synaptic activity and alterations in membrane potentials of neuron. These potentials are captured by surface EEG electrodes. But as presented in Fig. 1.6, there is also an intracellular current which flows from the synaptic area to the non-stimulated part of the neuron. The intracellular current is often called ***primary current***, while the extracellular - ***secondary, return or volume current***. Both primary and secondary currents contribute to magnetic fields outside the head [Hämäläinen et al., 1993, Baillet et al., 2001]. Fig. 1.10 shows the magnetic field generated by the primary current of an excited neuron.

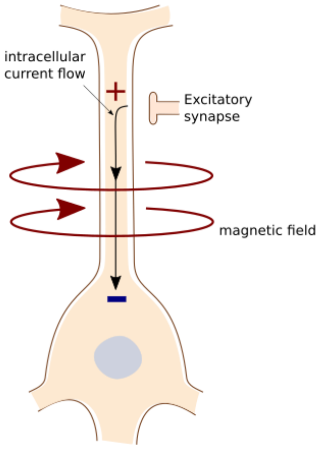


Figure 1.10: Schematic representation of a magnetic field generated by the primary current, resulting of EPSP.

The primary current is more dense than secondary current, as they integrate to the same magnitude, with the extracellular current being much more dispersed [Gloor, 1985]. However, in certain finite conductor geometries the volume current causes an equal but opposite field to that generated by the primary current. The net external field is then zero. For example, the radial sources of a spherically symmetric conductor are externally silent [Hämäläinen et al., 1993], which makes MEG sensitive mostly to the tangential sources.

1.7. From a single neuron to cortical activity

Calculations such as those in [Hämäläinen et al., 1993] suggest each synapse along a dendrite may contribute to as little as a 20 fA·m current source, too small to be measured with MEG/EEG. Empirical observations instead suggest that we are seeing contribution from sources on the order of 10 nA·m. In general terms, the electric and magnetic fields of a population of neurons equals the sum of the fields of the individual neurons. Hence, what we measure empirically is the cumulative summation of activity of millions of synapses in a relatively small region. Nominal calculations of neuronal density and cortical thickness suggest that the cortex has a macrocellular current density on the order of 100 nA/mm^2 . If we assume the cortex is about 4mm thick, then a small patch 5 mm × 5 mm would yield a net current of 10 nA-m, consistent with empirical observations and invasive studies [Hämäläinen et al., 1993, Baillet et al., 2001].

To produce a signal strong enough to be measured on the scalp or outside the head, the responsible neurons should be regularly arranged and activated in a more or less synchronous way [Niedermeyer and Da Silva, 2004]. If neurons have different orientations, the corresponding generated fields are likely to cancel each other out. If the neuron activity is not synchronized, the peaks of amplitude are not overlapping and thus the amplitude of the cumulative signal might not be high enough to be captured at a long distance.

Macrocolumns of tens of thousands of synchronously activated large ***pyramidal*** cortical neurons are thus believed to be the main MEG and EEG generators because of the coherent distribution of their large dendritic trunks locally oriented in parallel, and pointing perpendicularly to the cortical surface (Fig.1.11) [Baillet et al., 2001].

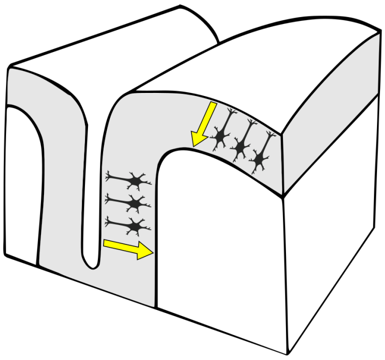


Figure 1.11: Schematic representation of population of neurons (black) with the main axis oriented perpendicular to the cortical surface. A dipole (yellow) is a reasonable electrical model for such a population. Note that the current flow may be of opposite direction, depending on the type of postsynaptic current (excitatory/inhibitory) and the locations of synapses. Adapted from [Hari et al., 2017].

1.8. EEG vs. MEG

Even though EEG and MEG are two modalities that measure the signals originated from the same physiological process - postsynaptic potentials, they are based on two different but complementary physical phenomena - electric and magnetic fields. As a result, they can provide complementary information about underlying neuronal activity (Fig.1.12).

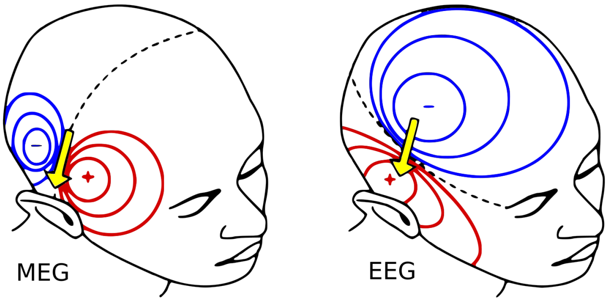


Figure 1.12: Schematic representation of the electrical potentials and magnetic field produced by the same dipole (yellow arrow) and measured on the scalp surface. Adapted form [Hari et al., 2017].

EEG is most sensitive to radial currents in radial parts of the cortex just under the electrode, but it can also sense tangential currents and strong deep currents. However, EEG has a limited spatial resolution, which is mainly due to the low conductivity of the skull, compared to that of brain and scalp, which makes the skull act like a spatial low-pass spatial filter of the signal [Hari et al., 2017]. It means that it may be difficult to distinguish multiple active sources from recorded signals.

The high sensitivity of MEG to tangential currents means that Mit mainly measures activity occurring on the walls of cortical fissures. This is an advantage, as about two-thirds of the cerebral cortex is located within fissures that are difficult places to reach even with intracranical recordings [Hari et al., 2017]. MEG is also less sensitive to conductivities of different head-tissue layers, which means less assumptions about forward model are required compared to EEG. However, as mentioned above, radial sources are invisible to MEG in a spherical head model. Even though the head is not a perfect sphere, MEG has a low sensitivity to radial – and consequently to deep – sources (any direction in the center of a sphere is radial).

All these factors mean that MEG and EEG have their own characteristics that make each of them an appropriate tool for some purposes but not for others [Hari et al., 2017]. The complementary nature of EEG and EEG signals also means that using both modalities sometimes provides information not obtainable with either modality alone [Yoshinaga et al., 2002, Dubarry et al., 2014, Baillet et al., 1999].

1.9. Conclusions

This chapter introduced the basic knowledge of the neurophysiological origins of the signals measured by EEG and MEG. This knowledge is crucial for understanding M/EEG forward and inverse problems, which we will discuss in the rest of this document. To summarize this chapter, we highlight several "take-home" messages:

- M/EEG signals are mostly generated by postsynaptic activity of macro-columns of tens of thousands of pyramidal neurons in the cortex.
- To generate a signal strong enough to be measured outside the brain, these neurons should be spatially aligned and temporally synchronized.
- The polarity of the signal depends on two factors, the type of the activated synapse (excitatory or inhibitory) and the position of the synapse along the postsynaptic cell body.
- The electromagnetic activity of local populations of synchronized and aligned neurons is often modeled as a dipole.
- EEG and MEG are complementary modalities.

Bibliography

Sylvain Baillet, Line Garnero, Gildas Marin, and Jean Paul Hugonin. Combined MEG and EEG source imaging by minimization of mutual information. *IEEE Transactions on Biomedical Engineering*, 46(5):522–534, 1999. ISSN 00189294. doi: 10.1109/10.759053.

Sylvain Baillet, John C. Mosher, and Richard M. Leahy. Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, 2001. ISSN 10535888. doi: 10.1109/79.962275.

Elena Boto, Sofie S. Meyer, Vishal Shah, Orang Alem, Svenja Knappe, Peter Kruger, T. Mark Fromhold, Mark Lim, Paul M. Glover, Peter G. Morris, Richard Bowtell, Gareth R. Barnes, and Matthew J. Brookes. A new generation of magnetoencephalography: Room temperature measurements using optically-pumped magnetometers. *NeuroImage*, 2017. ISSN 10959572. doi: 10.1016/j.neuroimage.2017.01.034.

György Buzsáki, Costas A. Anastassiou, and Christof Koch. The origin of extracellular fields and currents-EEG, ECoG, LFP and spikes, 2012. ISSN 1471003X.

Anne Sophie Dubarry, Jean Michel Badier, Agnès Trébuchon-Da Fonseca, Martine Gavaret, Romain Carron, Fabrice Bartolomei, Catherine Liégeois-Chauvel, Jean Régis, Patrick Chauvel, F. Xavier Alario, and Christian G. Bénar. Simultaneous recording of MEG, EEG and intracerebral EEG during visual stimulation: From feasibility to single-trial analysis. *NeuroImage*, 2014. ISSN 10959572. doi: 10.1016/j.neuroimage.2014.05.055.

P Gloor. Neuronal generators and the problem of localization in electroencephalography: application of volume conductor theory to electroencephalography. *Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society*, 2(4):327–354, oct 1985. ISSN 0736-0258 (Print).

Matti Hämäläinen, Riitta Hari, Risto J. Ilmoniemi, Jukka Knuutila, and Olli V. Lounasmaa. Magnetoencephalography theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, 1993. ISSN 00346861. doi: 10.1103/RevModPhys.65.413.

R. Hari, A. Puce, and Oxford University Press. *MEG-EEG Primer*. Oxford University Press, 2017. ISBN 9780190497804. URL `https://books.google.fr/books?id=duXOuQEACAAJ`.

Eric R. Kandel, James H. Schwartz, Thomas M. Jessel, Steven A. Siegelbaum, and A. J. Hudspeth. *Principles of neural science*. 5 edition, 2012. ISBN 9780071390118.

Henrik Lindén, Klas H. Pettersen, and Gaute T. Einevoll. Intrinsic dendritic filtering gives low-pass power spectra of local field potentials. *Journal of Computational Neuroscience*, 2010. ISSN 09295313. doi: 10.1007/s10827-010-0245-4.

Ernst Niedermeyer and F.H.Lopes Da Silva. *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, Fifth Edition*. Lippincott Williams and Wilkins, fifth edition, 2004. ISBN 0-7817-5126-8.

Harumi Yoshinaga, Tomoyuki Nakahori, Yoko Ohtsuka, Eiji Oka, Yoshihiro Kitamura, Hideki Kiriyama, Kazumasa Kinugasa, Keiichi Miyamoto, and Toru Hoshida. Benefit of simultaneous recording of EEG and MEG in dipole localization. *Epilepsia*, 2002. ISSN 00139580. doi: 10.1046/j.1528-1157.2002.42901.x.

Chapter 2

Forward modeling of EEG and MEG signals

Contents

2.1. Overview

In the previous chapter, we discussed the neurophysiological origins of the EEG and MEG signals. We introduced the electric and magnetic fields generated by neuronal activity but discussed them in a descriptive manner. In this chapter, we will introduce more explicit mathematical modeling of these processes. Starting with Maxwell's equations, we will introduce its quasi-static approximation which yields the Poisson equation with Neumann boundary condition. We will then review analytical and numerical methods for solving this equation.

2.2. Quasi-static approximation of Maxwell's equations

$$\nabla \cdot \boldsymbol{E} = \frac{\rho}{\epsilon_0} \tag{2.1}$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \tag{2.2}$$

$$\nabla \cdot \boldsymbol{B} = 0 \tag{2.3}$$

$$\nabla \times \boldsymbol{B} = \mu_0(\boldsymbol{J} + \epsilon_0 \frac{\partial \boldsymbol{E}}{\partial t}). \tag{2.4}$$

The useful frequency spectrum for electrophysiological signals in MEG and EEG is typically below 1 kHz, and most studies deal with frequencies between 0.1 and 100 Hz. Consequently, the physics of MEG and EEG can be described by the quasi-static approximation of Maxwell's equations [Baillet et al., 2001]. It means that terms $\frac{\partial \boldsymbol{B}}{\partial t}$ and $\epsilon_0 \frac{\partial \boldsymbol{E}}{\partial t}$ can be ignored as source terms [Hämäläinen et al., 1993]. Thus $\nabla \times \boldsymbol{E} = 0$ which means that $\boldsymbol{E}$ is an irrotational vector field. This also means that the field is conservative, which allows us to express the electric field as deriving from a scalar potential [Feynman, 2010].

$$\boldsymbol{E} = -\nabla V.$$

It is useful to divide the current density $\boldsymbol{J}$ produced by neuronal activity into two components, the volume extracellular secondary current $\boldsymbol{J}^v = \sigma \boldsymbol{E}$ and the intracellular primary current $\boldsymbol{J}^p$:

$$\boldsymbol{J} = \boldsymbol{J}^p + \sigma \boldsymbol{E} = \boldsymbol{J}^p - \sigma \nabla V. \tag{2.5}$$

Taking the divergence of both sides of Eq. (2.5), we get:

$$\nabla \cdot \boldsymbol{J} + \nabla \cdot (\sigma \nabla V) = \nabla \cdot \boldsymbol{J}^p.$$

The current density is divergence free in the quasistatic approximation, i.e. $\nabla \cdot \boldsymbol{J} = 0$. Indeed, from Eq. (2.4), since the divergence of a curl vanishes and $\epsilon_0 \frac{\partial \boldsymbol{E}}{\partial t}$ is ignored, we can write:

$$\begin{aligned}
&\nabla \cdot (\nabla \times \boldsymbol{B}) = \nabla \cdot \mu_0 (\boldsymbol{J} + \epsilon_0 \frac{\partial \boldsymbol{E}}{\partial t}) \\
&0 = \mu_0 \nabla \cdot \boldsymbol{J} + 0 \\
&\nabla \cdot \boldsymbol{J} = 0.
\end{aligned}$$

This shows that the electric potential and the primary current are related by a so-called *Poisson* equation:

$$\nabla \cdot (\sigma \nabla V) = \nabla \cdot \boldsymbol{J}^p. \tag{2.6}$$

The Biot and Savart law relates the magnetic field to the current density:

$$\boldsymbol{B}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} \int \boldsymbol{J}(\boldsymbol{r'}) \times \frac{\boldsymbol{r} - \boldsymbol{r'}}{\|\boldsymbol{r} - \boldsymbol{r'}\|^3} d\boldsymbol{r'}.$$

Using the decomposition of current density $\boldsymbol{J}$ from Eq. (2.5), we get:

$$\begin{aligned}
\boldsymbol{B}(\boldsymbol{r}) &= \frac{\mu_0}{4\pi} \int \boldsymbol{J}^p(\boldsymbol{r'}) \times \frac{\boldsymbol{r} - \boldsymbol{r'}}{\|\boldsymbol{r} - \boldsymbol{r'}\|^3} d\boldsymbol{r'} - \frac{\mu_0}{4\pi} \int \sigma(\boldsymbol{r'}) \nabla V(\boldsymbol{r'}) \times \frac{\boldsymbol{r} - \boldsymbol{r'}}{\|\boldsymbol{r} - \boldsymbol{r'}\|^3} d\boldsymbol{r'} \\
&= \boldsymbol{B}_0 + \boldsymbol{B}_{sec},
\end{aligned} \tag{2.7}$$

where $\boldsymbol{B}_0$ is the contribution to the magnetic field coming from the primary current and $\boldsymbol{B}_{sec}$ is the contribution of volume (secondary) currents.

2.3. Source modeling

The most common source model to represent electrical activity in the brain is a "current dipole". It represents an oriented source of current located at a single position $\boldsymbol{r}_0$, with dipolar moment $\boldsymbol{q}$, and it is denoted by:

$$J^p(\boldsymbol{r}) = \boldsymbol{q} \delta(\boldsymbol{r} - \boldsymbol{r}_0), \tag{2.8}$$

where $\delta(\cdot)$ is a Dirac delta distribution.

2.4. Head modeling

As we can see from Eq. (2.6) and Eq. (2.7), the conductor model (or also **head model**) is required to define these equations, i.e. the conductivity distribution

$\sigma(\boldsymbol{r})$ must be specified. In this section we will discuss different approaches with increasing model complexities.

2.4.1. Infinite homogeneous medium

The main property of the homogeneous medium is that the conductivity is constant, and so:

$$\nabla \cdot (\sigma \nabla V) = \sigma \Delta V.$$

Assuming the dipole model for the primary current at location $\boldsymbol{p}$ (see Eq. (2.8)), Eq. (2.6) becomes:

$$\Delta V = \frac{1}{\sigma} \boldsymbol{q} \cdot \nabla \delta(\boldsymbol{r} - \boldsymbol{p}) \tag{2.9}$$

This is a Poisson equation, which general form is $\Delta u = f$ and whose particular solution is:

$$u(\boldsymbol{r}) = \int_{\mathbb{R}^3} G_L(\boldsymbol{r} - \boldsymbol{r'}) f(\boldsymbol{r'}) d\boldsymbol{r'},$$

where $G_L(\boldsymbol{r}) = -\dfrac{1}{4\pi\|\boldsymbol{r}\|}$ is the Green function of the Laplacian.

Thus, the solution of Eq. (2.9) is:

$$V(\boldsymbol{r}) = -\frac{1}{4\pi\sigma} \int \boldsymbol{q} \cdot \nabla \delta(\boldsymbol{r'} - \boldsymbol{p}) \frac{1}{\|\boldsymbol{r} - \boldsymbol{r'}\|} d\boldsymbol{r'} = \frac{1}{4\pi\sigma} \boldsymbol{q} \cdot \frac{\boldsymbol{r} - \boldsymbol{p}}{\|\boldsymbol{r} - \boldsymbol{p}\|^3}.$$

In an infinite, homogeneous domain, it can be shown that only the primary current contributes to the magnetic field [Hämäläinen et al., 1993], therefore, for a dipolar source:

$$\boldsymbol{B}(\boldsymbol{r}) = \boldsymbol{B}_0(\boldsymbol{r}) = \frac{\mu_0}{4\pi} \boldsymbol{q} \times \frac{\boldsymbol{r} - \boldsymbol{p}}{\|\boldsymbol{r} - \boldsymbol{p}\|^3}.$$

2.4.2. Piecewise homogeneous conductor

An infinite, homogeneous domain is not a reasonable model for actual M/EEG signals measured on the scalp or outside the head. First of all, air does not conduct electricity and so the conductivity outside the head should equal to zero. Moreover, the head consists of different tissues and is not homogeneous at all in terms of conductivity.

Let us first start again with the Poisson equation (2.6), constrained but the boundary condition, which is the hypothesis that no current flows outside of the head. This is mostly true except at the neck which is far from most of the M/EEG measurements, so we can assume that it has a negligible effect. Let $\Omega \subset \mathbb{R}^3$ denote the head domain, $\partial\Omega$ its boundary (scalp surface) with outward pointing normal vector $\boldsymbol{n}$. We thus have to solve the following forward problem:

$$\begin{cases} \nabla \cdot (\sigma \nabla V) = f = \nabla \cdot \boldsymbol{J}^p & \text{in } \Omega \\ \sigma \dfrac{\partial V}{\partial \boldsymbol{n}} = \sigma \nabla V \cdot \boldsymbol{n} = 0 & \text{on } \partial\Omega \ . \end{cases} \tag{2.10}$$

The second line of the equation (boundary condition) reflects the hypothesis that no current flows out of the head.

Several nonhomogeneous head models are used in practice. A first refinement is the case where the head is assumed to comprise a set of *nested concentric spheres*, each of constant conductivity. In this case, analytic solutions exist for both MEG and EEG [Mosher et al., 1999]. Analytic solutions for other head shapes have been presented, such as prolate and oblate spheroids [Cuffin and Cohen, 1977] or eccentric spheres [Cuffin, 1991]. Numerical solutions for narrow or wide ellipsoids are presented in [Cuffin, 1990].

In the case of a concentric spherical head model, the radial component of the magnetic field depends only on the primary current and is independent of the conductivity profile. Moreover, even though the tangential field components are affected by the volume currents, they also can be computed without knowing the conductivity profile [Hämäläinen et al., 1993].

Another important point to note is that a radial source in a spherical model will produce no magnetic field outside the head [Hämäläinen et al., 1993, Grynszpan and Geselowitz, 1973].

A head model is **semi-realistic** when the regions of the piecewise homogeneous conductor are arbitrarily shaped (Fig. 2.1). In this case, the solutions of M/EEG forward problem must be computed numerically. We will discuss these numerical methods in the Section 2.5.

2.4.3. Realistic head model

For even more realistic models, the piecewise homogeneity of conductivity that has been made in the previous section needs to be relaxed. Indeed, the brain is known to have strong anisotropies in the conductivities at least in two domains: the skull and the white matter (Fig. 2.2).

The human skull consists of a soft bone layer (spongiosa) enclosed by two hard bone layers (compacta). As the spongiosa has a much higher conductivity than the compacta [Akhtari et al., 2002], the skull can be described by an effective anisotropic conductivity with a ratio of up to 1:10 radially to tangentially to the skull surface [Rush and Driscoll, 1968, Wolters et al., 2006].

It is also known that brain white matter – part of the brain that is mainly made up of myelinated axons – has an anisotropic conductivity with a ratio of

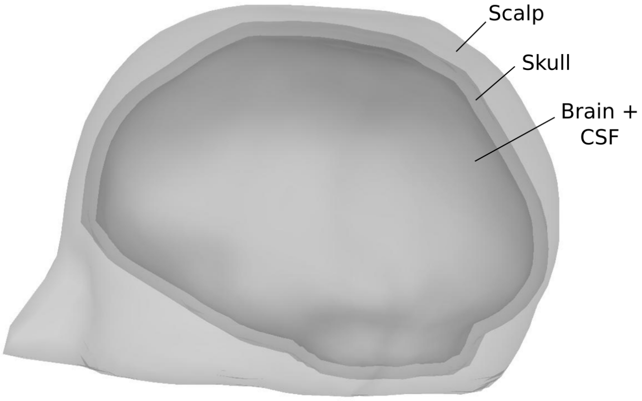


Figure 2.1: Example of a piecewise homogeneous conductor model of the head. Three layers are used: Scalp, skull and brain and cerebrospinal fluid (CSF).

about 1:10 (normal:parallel to fibers) [Nicholson, 1965, Wolters et al., 2006] and several methods have been developed to use diffusion tensor imaging for modelling this white matter conductivity anisotropy [Güllmar et al., 2010, Lee et al., 2009]. In this case the conductivity is modeled not as a scalar but as a three-dimensional tensor, associated to each point of space.

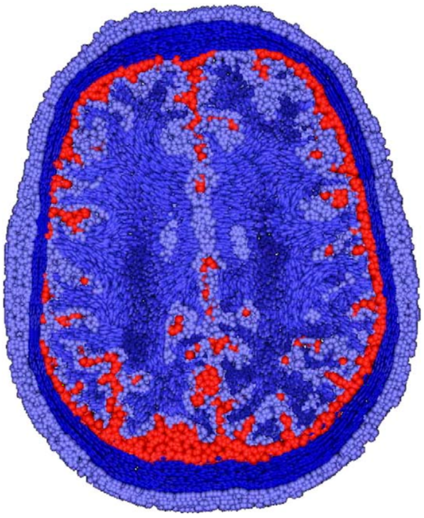


Figure 2.2: Conductivity tensor ellipsoids in the barycenters of the tetrahedra elements: Normalized and colored for skull and white matter anisotropy. The highest trace values can be found in the CSF compartment (red) and the lowest in the skull compartment (dark blue). Taken from [Wolters et al., 2006].

2.5. Numerical methods to solve M/EEG forward problem

As mentioned in the previous sections, for semi-realistic and realistic head models the solution of the M/EEG forward problem does not exist in a closed form, and thus numerical methods are required to solve it. In this section, we will focus on the finite element method (FEM) [Wolters et al., 2004, Vallaghé and Papadopoulo, 2010] and the boundary element method (BEM) [Sarvas, 1987, Kybic et al., 2005].

2.5.1. Finite element method (FEM)

Let us first define some functional spaces that will be needed hereafter:

$$H^1(\Omega) = \{w \in L^2(\Omega), \nabla w \in L^2(\Omega)^3\},$$

$$H^2(\Omega) = \{w \in L^2(\Omega), \nabla w \in H^1(\Omega)^3\}.$$

1. *Variational formulation.*

   It can be proven that the following optimization problem is equivalent to Eq. (2.10) for $V \in H^2(\Omega)$:

$$\begin{aligned} &V = \underset{\phi \in H^1(\Omega)}{\operatorname{argmin}} E(\phi) \text{ where:} \\ &E(\phi) = \int_\Omega \sigma(\boldsymbol{r}) \left\|\nabla\phi(\boldsymbol{r})\right\|^2 d\boldsymbol{r} + \int_\Omega f(\boldsymbol{r}) w(\boldsymbol{r}) d\boldsymbol{r} \end{aligned} \tag{2.11}$$

   This is so-called variational formulation of the initial problem (2.10).

2. *Discretization.*

   The 3D domain Ω is tessellated with bounded cells $C_i,\ i = 1...N_C$. This also introduces a set of vertices $V_i,\ i = 1...N_V$. The space of continuous functions over Ω is approximated by a finite-dimensional vector space which is a spanned by a set of basis functions $w^i, i = 1...N_V$:

$$H^1_h(\Omega_h) = \operatorname{span}(\{w^i(\boldsymbol{r}), i = 1...N_V\})$$

   Notice that that σ is assumed to be constant within each cell.

3. *Galerkin projection.*

We can approximate $E(\phi)$ from Eq. (2.11) by its discretized version $E_h(\boldsymbol{v})$, with $\phi(\boldsymbol{r}) \approx \sum_{i=1}^{N_V} v_i w^i(\boldsymbol{r})$:

$$E_h(\boldsymbol{v}) = \frac{1}{2}\sum_{i,j=1}^{N_V} v_i v_j \int_\Omega \sigma(\boldsymbol{r})\nabla w^i(\boldsymbol{r})\nabla w^j(\boldsymbol{r})d\boldsymbol{r} + \sum_{i=1}^{N_V} v_i \int_\Omega f(\boldsymbol{r})w^i(\boldsymbol{r})d\boldsymbol{r} - \sum_{i=1}^{N_S} v_i \int_S g(\boldsymbol{r})w^i_{S_h}(\boldsymbol{r})ds \tag{2.12}$$

Now the variable is no longer a continuous function but a finite vector $\boldsymbol{v} = (v_1, ..., v_{N_V}) \in \mathbb{R}^{N_V}$. Let us also notice that in our case $g \equiv 0$, so the last term in $E(\boldsymbol{v})$ vanishes.

Let us introduce the matrix $\boldsymbol{H}$ such that:

$$H_{ij} = \int_\Omega \sigma(\boldsymbol{r})\nabla w^i(\boldsymbol{r})\nabla w^j(\boldsymbol{r})d\boldsymbol{r} \tag{2.13}$$

and the vector $\boldsymbol{d}$:

$$d_i = -\int_\Omega f(\boldsymbol{r})w^i(\boldsymbol{r})d\boldsymbol{r}$$

We can then formulate the discrete version of the optimization problem (2.11):

$$V = \sum_{i}^{N_V} v_i w^i, \text{ where: } \boldsymbol{v} = \underset{\overline{\boldsymbol{v}}}{\operatorname{argmin}}\, E_h(\overline{\boldsymbol{v}}) \tag{2.14}$$

The criterion $E_h(\boldsymbol{v})$ can be written as $\frac{1}{2}\boldsymbol{v}^T\boldsymbol{H}\boldsymbol{v} + \boldsymbol{d}\boldsymbol{v}$ and thus the optimal $\boldsymbol{v}$ of Eq. (2.14) is the solution of the following linear system:

$$\boldsymbol{H}\boldsymbol{v} = \boldsymbol{d}. \tag{2.15}$$

Note that the so-called *head matrix* $\boldsymbol{H}$ is symmetric. Usually, the basis functions will have a very local support, so that $\boldsymbol{H}$ will be very sparse. It also has a one dimensional kernel and thus is not invertible.

Eq. (2.15) is a fairly simple matrix problem, which could in theory be solved using a pseudo-inverse. However, $\boldsymbol{H}$ is usually a very big matrix. In this case, it is better to solve the problem using an iterative method (a lot of efficient methods have been developed to do it).

Choice of basis functions

In practice, the basis of P^1 functions over a tetrahedral mesh is used. In dimension d, a tetrahedron T_j is defined by $d+1$ vertices $V_{i_k}, k = 1, ..., d+1$. The restriction w_j^i of w^i to T_j is defined by (we assume without loss of generality that $V_{i_k} = V_k$):

$$w_j^i(\boldsymbol{r}) = \begin{cases} \dfrac{|\boldsymbol{r}V_2...V_{d+1}|}{|V_1...V_{d+1}|} & \text{, if } \boldsymbol{r} \in T_j, \\ 0 & \text{, otherwise.} \end{cases}$$

In other words, space $H_h^1(\Omega_h)$ is the space of functions which are affine per tetrahedron. Moreover, the values of ϕ correspond to function values at vertices.

2.5.2. Symmetric boundary element method

The boundary element method (BEM) is based on integral equations involving unknowns at interfaces between domains Ω_i of homogeneous conductivity (Fig. 2.3), where FEM consider the entire volume. We also assume conductivity to be isotropic in each head subdomain.

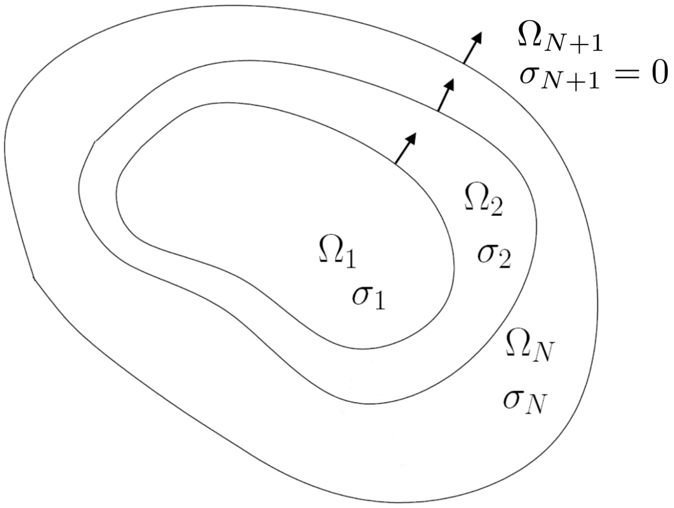


Figure 2.3: Without loss of generality the head is modeled as a set of regions $\Omega_1, ..., \Omega_N$ with constant isotropic conductivities $\sigma_1, ..., \sigma_N$. Adapted from Kybic et al. [2005].

We use the fundamental Representation Theorem of potential theory, which shows that a harmonic function u is determined everywhere in $\mathbb{R}^3$ by its jump and the jump of its derivative across a boundaries $\cup\partial\Omega_i$. For more details see Kybic et al. [2005].

We define the four integral operators:

$$(\mathcal{D}f)(\boldsymbol{r}) = -\int_{\cup\partial\Omega_i} \partial_{\boldsymbol{n}'} G(\boldsymbol{r}-\boldsymbol{r}')f(\boldsymbol{r}')ds(\boldsymbol{r}'),$$

$$(\mathcal{S}f)(\boldsymbol{r}) = -\int_{\cup\partial\Omega_i} G(\boldsymbol{r}-\boldsymbol{r}')f(\boldsymbol{r}')ds(\boldsymbol{r}'),$$

$$(\mathcal{N}f)(\boldsymbol{r}) = -\int_{\cup\partial\Omega_i} \partial^2_{\boldsymbol{n},\boldsymbol{n}'} G(\boldsymbol{r}-\boldsymbol{r}')f(\boldsymbol{r}')ds(\boldsymbol{r}'),$$

$$(\mathcal{D}^*f)(\boldsymbol{r}) = -\int_{\cup\partial\Omega_i} \partial_{\boldsymbol{n}} G(\boldsymbol{r}-\boldsymbol{r}')f(\boldsymbol{r}')ds(\boldsymbol{r}'),$$

where $\boldsymbol{r} \in \mathbb{R}^3$, $\boldsymbol{n}'$ denotes the outward normal vector at $\boldsymbol{r}' \in \cup\partial\Omega_i$ and $G(\boldsymbol{r}) = -\frac{1}{4\pi\|\boldsymbol{r}\|}$ is a Green function of the Laplacian.

In addition of potentials on the surfaces V_{S_i} which are unknown in our model, the symmetric BEM considers the normal current $p_{S_i} = \sigma_i(\partial_{\boldsymbol{n}}V)_{S_i}$ as additional unknowns.

So we have two sets of unknowns V_{S_i} and p_{S_i} which are continuous functions on the interfaces between domains and we can discretize them using a finite number of basis functions. We want the approximation errors of the two quantities (V_S and p_S) to be asymptotically equivalent. For this reason, we choose to approximate V using $P1$ basis functions as $V_{S_k}(\boldsymbol{r}) = \sum_i x_i^{(k)}\phi_i^{(k)}(\boldsymbol{r})$, while its normal derivative p is represented using the space $P0$, $p_{S_k}(\boldsymbol{r}) = \sum_i y_i^{(k)}\psi_i^{(k)}(\boldsymbol{r})$. $P0$ basis function ψ_i equals to 1 on mesh triangle T_i and 0 elsewhere.

After the discretization is done we can compute the unknown projection coefficients by solving the system (2.16).

$$\begin{bmatrix} (\sigma_1+\sigma_2)\boldsymbol{N}_{11} & -2\boldsymbol{D}^*_{11} & -\sigma_2\boldsymbol{N}_{12} & \boldsymbol{D}^*_{12} & & & \\ -2\boldsymbol{D}_{11} & (\sigma_1^{-1}+\sigma_2^{-1})\boldsymbol{S}_{11} & \boldsymbol{D}_{12} & -\sigma_2^{-1}\boldsymbol{S}_{12} & & & \\ -\sigma_2\boldsymbol{N}_{21} & \boldsymbol{D}^*_{21} & (\sigma_2+\sigma_3)\boldsymbol{N}_{22} & -2\boldsymbol{D}^*_{22} & -\sigma_3\boldsymbol{N}_{23} & \boldsymbol{D}^*_{23} & \\ \boldsymbol{D}_{21} & -\sigma_2^{-1}S_{21} & -2\boldsymbol{D}_{22} & (\sigma_2^{-1}+\sigma_3^{-1})\boldsymbol{S}_{22} & \boldsymbol{D}_{23} & -\sigma_3^{-1}S_{23} & \\ & & -\sigma_3\boldsymbol{N}_{32} & \boldsymbol{D}^*_{32} & (\sigma_3+\sigma_4)\boldsymbol{N}_{33} & -2\boldsymbol{D}^*_{33} & \dots \\ & & \boldsymbol{D}_{32} & -\sigma_3^{-1}\boldsymbol{S}_{32} & -2\boldsymbol{D}_{33} & (\sigma_3^{-1}+\sigma_4^{-1})\boldsymbol{S}_{33} & \dots \\ & & & & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{y}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{y}_2 \\ \boldsymbol{x}_3 \\ \boldsymbol{y}_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{c}_1 \\ \boldsymbol{b}_2 \\ \boldsymbol{c}_2 \\ \boldsymbol{b}_3 \\ \boldsymbol{c}_3 \\ \vdots \end{bmatrix} \tag{2.16}$$

$$(N_{kl})_{ij} = \langle \mathcal{N}_{kl}\phi_j^{(l)}, \phi_i^{(k)}\rangle \qquad (\boldsymbol{b}_k)_i = \langle \partial_{\boldsymbol{n}}v_{\Omega_k} - \partial_{\boldsymbol{n}}v_{\Omega_{k+1}},\ \phi_i^{(k)}\rangle$$

$$(D_{kl})_{ij} = (D^*_{lk})_{ji} = \langle \mathcal{D}_{kl}\phi_j^{(l)}, \psi_i^{(k)}\rangle \qquad (\boldsymbol{c}_k)_i = \langle \sigma_{k+1}^{-1}v_{\Omega_{k+1}} - \sigma_k^{-1}v_{\Omega_k},\ \psi_i^{(k)}\rangle$$

$$(S_{kl})_{ij} = \langle \mathcal{S}_{kl}\psi_j^{(l)}, \psi_i^{(k)}\rangle \qquad (\boldsymbol{x}_k)_i = x_i^{(k)},\ (\boldsymbol{y}_k)_i = y_i^{(k)}.$$

Here v_{Ω_i} represents the homogeneous medium potential $v_{\Omega_i}(\boldsymbol{r}) = f_{\Omega_i} * G(\boldsymbol{r})$, where $f_{\Omega_i} = f \cdot \mathbf{1}_{\Omega_i} = \nabla \cdot J^p \cdot \mathbf{1}_{\Omega_i}$.

2.5.3. Generalization

Let us notice that for both the FEM and the symmetric BEM, the initial continuous problem (Eq. 2.10) is discretized to a linear system of type:

$$\boldsymbol{H}\boldsymbol{v} = \boldsymbol{d}. \tag{2.17}$$

Let N_V be the number of unknowns in the head model. Vector $\boldsymbol{v} \in \mathbb{R}^{N_V}$ represents unknowns of the model, which are values of the potential on the mesh nodes for FEM and some standard BEM, and potentials and their normal derivatives on the meshes for symmetric BEM. The $N_V \times N_V$ matrix $\boldsymbol{H}$, called "head matrix", can be computed once the head geometry, its conductivity values and finite element basis functions are fixed. Vector $\boldsymbol{d}$ depends on the source configuration. Notice that $\boldsymbol{H} = \boldsymbol{H}_{\boldsymbol{\sigma}}$ depends on conductivity. An important property of matrix $\boldsymbol{H}_{\boldsymbol{\sigma}}$ is that, under the assumption that the head is composed of several subdomains with constant conductivity, it can be represented as a linear combination of conductivity-independent matrices:

$$\boldsymbol{H}_{\boldsymbol{\sigma}} = \sum_{i=1}^{N_H} \gamma_i(\boldsymbol{\sigma}) \overline{\boldsymbol{H}}_i, \tag{2.18}$$

where $\gamma_i(\boldsymbol{\sigma})$ are scalar functions, N_H represents the number of conductivity-independent components $\overline{\boldsymbol{H}}_i$. For FEM, $\gamma_i(\boldsymbol{\sigma}) = \sigma_i$ (from Eq. (2.13)). For symmetric BEM, multipliers $\gamma_i(\boldsymbol{\sigma})$ have more complex structure, for example $\{-\sigma_i,\ \sigma_i^{-1},\ \sigma_i + \sigma_j,\ \sigma_i^{-1} + \sigma_j^{-1},\ \cdots\}$ (from Eq. (2.16)). As an example, for a three-layer nested head model with brain, skull and scalp conductivities, $N_H = 3$ for FEM and $N_H = 7$ for BEM.

2.6. Source space modeling

The source space can be seen as a finite set of N_S dipoles with known positions. If the dipole orientation is known, a dipole is characterized by one amplitude. If not, the dipole orientation can be decomposed into a linear combination of three 3d unit basis vectors, and a dipole is characterized by three amplitudes. Without loss of generality, we can assume that the dipole orientation is known. Let x_i represent the amplitude of i-th source and $\bar{J}_i^p$ the i-th source with unit amplitude:

$$J^p(\boldsymbol{r}) = \sum_{i=1}^{N_S} x_i \bar{J}_i^p(\boldsymbol{r}).$$

The source term $\boldsymbol{d}$ in Eq. (2.17) is linear with respect to J, therefore:

$$\boldsymbol{d}(J^p) = \sum_{i=1}^{N_S} x_i \boldsymbol{d}(\bar{J}_i^p(\boldsymbol{r})) = \sum_{i=1}^{N_S} x_i \boldsymbol{d}_i = \boldsymbol{D}\boldsymbol{x} \ ,$$

where the i-th column of the $N_V \times N_S$ matrix $\boldsymbol{D}$ corresponds to the i-th unit source and $\boldsymbol{x} = (x_1, \cdots, x_{N_S})$ is the vector of source amplitudes. Let us notice that in the case of FEM, matrix $\boldsymbol{D}$ does not depend on conductivities. For BEM, it does, so we will note it $\boldsymbol{D}_{\boldsymbol{\sigma}}$ for generality. $\boldsymbol{D}_{\boldsymbol{\sigma}}$ can be decomposed as:

$$\boldsymbol{D}_{\boldsymbol{\sigma}} = \sum_{i=1}^{N_D} \lambda_i(\boldsymbol{\sigma}) \overline{\boldsymbol{D}}_i \ , \tag{2.19}$$

where matrices $\overline{\boldsymbol{D}}_i$ are independent of $\boldsymbol{\sigma}$ and $\lambda_i(\boldsymbol{\sigma})$ are scalars. In the case of symmetric BEM $\lambda_i(\boldsymbol{\sigma})$ are $\{1, \sigma_i^{-1}, \cdots\}$ (from Eq. (2.16)).

The linear system to solve then becomes:

$$\boldsymbol{H}_{\boldsymbol{\sigma}}\boldsymbol{v} = \boldsymbol{D}_{\boldsymbol{\sigma}}\boldsymbol{x}. \tag{2.20}$$

2.7. M/EEG lead field matrix

Usually, computing the M/EEG forward problem means computing the **lead field matrix**. In the context of EEG, the lead field is the linear operator which maps source activation to potentials at sensor locations:

$$\boldsymbol{v}_{eeg} = \boldsymbol{L}_{eeg}\boldsymbol{x}.$$

Computing $\boldsymbol{v}_{eeg}$ requires applying to $\boldsymbol{v}$ (solution of the system (2.20)) a matrix S_{eeg} which selects or interpolates potentials only at electrode positions: $\boldsymbol{v}_{eeg} = \boldsymbol{S}_{eeg}\boldsymbol{v}$. Let N_E be the number of EEG electrodes.

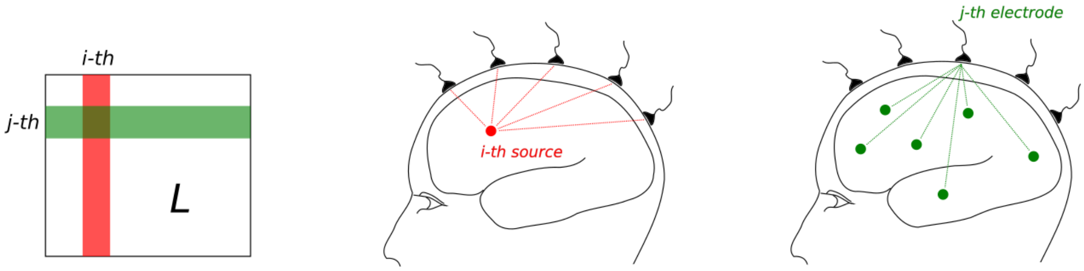


Figure 2.4: The i-th column of the lead field matrix $\boldsymbol{L}$ shows how the i-th source is seen by all electrodes. The j-th raw of the lead field shows how each source is seen by the j-th electrode.

Using Eq. (2.20) and the selection matrix $\boldsymbol{S}_{eeg}$:

$$\boldsymbol{v}_{eeg} = \boldsymbol{S}_{eeg}\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1}\boldsymbol{D}_{\boldsymbol{\sigma}}\boldsymbol{x}.$$

Let us remark that since the electric potential is only defined up to a constant, the head matrix $\boldsymbol{H}_{\boldsymbol{\sigma}}$ is not full rank and has a one-dimensional kernel. So the inverse notation $\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1}$ actually implies a deflation [Chan, 1984] which is usually applied to this type of situation.

The $N_E \times N_S$ lead field matrix can thus be computed as:

$$\boldsymbol{L}_{eeg} = \boldsymbol{S}_{eeg}\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1}\boldsymbol{D}_{\boldsymbol{\sigma}}. \tag{2.21}$$

Each column of the lead field matrix represents the contribution of the corresponding unit norm source on the EEG electrodes (Fig. 2.4).

In the context of MEG, the lead field is a linear operator that maps source activation to magnetic measurements. A magnetometer measures the flux of the magnetic field through a small loop. For a magnetometer i, let M_i be the surface enclosed by the loop, and $\boldsymbol{d}_i$ the unitary vector normal to M_i. Thus the signal measured by a magnetometer is modeled as $\int_{M_i} \boldsymbol{d}_i \cdot \boldsymbol{B}(\boldsymbol{r}_i) d\boldsymbol{r}_i$.

Using magnetic field decomposition (Eq. (2.7)), the MEG lead field can be represented as:

$$\boldsymbol{L}_{meg} = \boldsymbol{L}_0 + \boldsymbol{S}_{meg}\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1}\boldsymbol{D}_{\boldsymbol{\sigma}}$$

where first component and second components correspond to the primary and secondary currents respectively. $\boldsymbol{S}_{meg}$ is a linear operator which maps electric potentials to magnetic flux at coils. Let us remark that $\boldsymbol{S}_{meg}$ is more complex than just a selection matrix $\boldsymbol{S}_{eeg}$. $\boldsymbol{L}_0$ can be computed analytically while $\boldsymbol{S}_{meg}$ is assembled based on the basis functions used for Galerkin projection. See [Wolters et al., 2004] for details.

2.8. Conclusions

This chapter introduced basic knowledge the about forward modeling of the M/EEG signals. To summarize this chapter, we highlight some important points:

- For realistic head models, numerical methods are used to solve the M/EEG forward problem.
- FEM is based on the discretization of the head volume, while the BEM only uses interfaces between different head domains.
- Both FEM and BEM require solving a linear system of type Eq. 2.17.
- In this work, computing the M/EEG forward problem will mean computing the lead field matrix (Eq. 2.21).
- Solving the M/EEG forward problem requires assembling matrices $\boldsymbol{S}, \boldsymbol{H_\sigma}$ and $\boldsymbol{D_\sigma}$, and computing $\boldsymbol{S}\boldsymbol{H_\sigma}^{-1}$.
- Matrices $\boldsymbol{H_\sigma}$ and $\boldsymbol{D_\sigma}$ can be expressed as a linear combination of conductivity independent matrices (Eq. 2.18 and Eq. 2.19).

Bibliography

M. Akhtari, H. C. Bryant, A. N. Mamelak, E. R. Flynn, L. Heller, J. J. Shih, M. Mandelkem, A. Matlachov, D. M. Ranken, E. D. Best, M. A. Dimauro, R. R. Lee, and W. W. Sutherling. Conductivities of three-layer line human skull. *Brain Topography*, 14(3):151–167, 2002. ISSN 08960267. doi: 10.1023/A:1014590923185.

Sylvain Baillet, John C. Mosher, and Richard M. Leahy. Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, 2001. ISSN 10535888. doi: 10.1109/79.962275.

Tony F. Chan. Deflated decomposition of solutions of nearly singular systems. *SIAM Journal on Numerical Analysis*, 21(4):738–754, 1984. ISSN 00361429. URL `http://www.jstor.org/stable/2157006`.

B. Neil Cuffin. Effects of Head Shape on EEG's and MEG's. *IEEE Transactions on Biomedical Engineering*, 37(1):44–52, 1990. ISSN 15582531. doi: 10.1109/10.43614.

B. Neil Cuffin. Eccentric Spheres Models of the Head. *IEEE Transactions on Biomedical Engineering*, 38(9):871–878, 1991. ISSN 15582531. doi: 10.1109/10.83607.

B. Neil Cuffin and David Cohen. Magnetic Fields of a Dipole in Special Volume Conductor Shapes. *IEEE Transactions on Biomedical Engineering*, BME-24(4):372–381, 1977. ISSN 15582531. doi: 10.1109/TBME.1977.326145.

Richard Feynman. *The Feynman lectures on physics*, volume 2. Basic Books, New York, 2010. ISBN 0-465-02416-5.

Flavio Grynszpan and David B. Geselowitz. Model Studies of the Magnetocardiogram. *Biophysical Journal*, 13(9):911–925, 1973. ISSN 00063495. doi: 10.1016/S0006-3495(73)86034-5.

Daniel Güllmar, Jens Haueisen, and Jürgen R Reichenbach. Influence of anisotropic electrical conductivity in white matter tissue on the EEG/MEG forward and inverse solution. A high-resolution whole head simulation study. *NeuroImage*, 51(1):145–163, 2010. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2010.02.014. URL `http://www.sciencedirect.com/science/article/pii/S1053811910001825`.

Matti Hämäläinen, Riitta Hari, Risto J. Ilmoniemi, Jukka Knuutila, and Olli V. Lounasmaa. Magnetoencephalography theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, 1993. ISSN 00346861. doi: 10.1103/RevModPhys.65.413.

Jan Kybic, Maureen Clerc, Toufic Abboud, Olivier Faugeras, Renaud Keriven, and Théo Papadopoulo. A common formalism for the Integral formulations of the forward EEG problem. *IEEE Transactions on Medical Imaging*, 24(1):12–28, jan 2005. ISSN 0278-0062. doi: 10.1109/TMI.2004.837363. URL `http://ieeexplore.ieee.org/document/1375158/`.

Won Hee Lee, Zhongming Liu, Bryon A. Mueller, Kelvin Lim, and Bin He. Influence of white matter anisotropic conductivity on EEG source localization: Comparison to fMRI in human primary visual cortex. *Clinical Neurophysiology*, 120(12):2071–2081, dec 2009. ISSN 13882457. doi: 10.1016/j.clinph.2009.09.007. URL `https://linkinghub.elsevier.com/retrieve/pii/S1388245709005355`.

John C. Mosher, Richard M. Leahy, and Paul S. Lewis. EEG and MEG: Forward solutions for inverse methods. *IEEE Transactions on Biomedical Engineering*, 46(3):245–259, 1999. ISSN 00189294. doi: 10.1109/10.748978.

Paul W. Nicholson. Specific impedance of cerebral white matter. *Experimental Neurology*, 13(4):386–401, 1965. ISSN 10902430. doi: 10.1016/0014-4886(65)90126-3.

S Rush and D A Driscoll. Current distribution in the brain from surface electrodes. *Anesthesia and analgesia*, 47(6):717–23, 1968. ISSN 0003-2999. URL `http://www.ncbi.nlm.nih.gov/pubmed/4972743`.

J. Sarvas. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Physics in Medicine and Biology*, 1987. ISSN 00319155. doi: 10.1088/0031-9155/32/1/004.

Sylvain Vallaghé and Théodore Papadopoulo. A Trilinear Immersed Finite Element Method for Solving the Electroencephalography Forward Problem. *SIAM Journal on Scientific Computing*, 32(4):2379–2394, jan 2010. ISSN 1064-8275. doi: 10.1137/09075038X. URL `http://epubs.siam.org/doi/10.1137/09075038X`.

C. H. Wolters, L. Grasedyck, and W. Hackbusch. Efficient computation of lead field bases and influence matrix for the FEM-based EEG and MEG inverse problem. *Inverse Problems*, 2004. ISSN 02665611. doi: 10.1088/0266-5611/20/4/007.

C. H. Wolters, A. Anwander, X. Tricoche, D. Weinstein, M. A. Koch, and R. S. MacLeod. Influence of tissue conductivity anisotropy on EEG/MEG field and return current computation in a realistic head model: A simulation and visualization study using high-resolution finite element modeling. *NeuroImage*, 30(3): 813–826, 2006. ISSN 10538119. doi: 10.1016/j.neuroimage.2005.10.014.

Chapter 3
M/EEG inverse problem

Contents

3.1. Overview

The M/EEG inverse problem consists in reconstruction and localization of the cortical sources activity, given a set of measurements and forward (lead field) operator. In this chapter, we will give an overview of several classes of state-of-art M/EEG source localization methods. Starting with the general formulation of well- and ill-posed problems, we will introduce basic principles of dipole fit and scanning methods, as well as variational and Bayesian approaches to solve the M/EEG inverse problem.

We assume that source orientations are known and fixed, unless otherwise specified. The forward model for a single time sample with fixed source orientations is:

$$\boldsymbol{y} = \sum_i \boldsymbol{L}_i x_i + \boldsymbol{b} = \boldsymbol{L}\boldsymbol{x} + \boldsymbol{b}, \tag{3.1}$$

where $\boldsymbol{y} \in \mathbb{R}^{N_E}$ is a measurements vector, $\boldsymbol{b} \in \mathbb{R}^{N_E}$ is additive noise, $\boldsymbol{x} \in \mathbb{R}^{N_S}$ is source amplitudes vector and $\boldsymbol{L}$ is a $N_E \times N_S$ lead field matrix.

In the case of N_T time samples, we will use the following notation:

$$\boldsymbol{Y} = \boldsymbol{L}\boldsymbol{X} + \boldsymbol{B}, \tag{3.2}$$

where $\boldsymbol{Y}$ and $\boldsymbol{B}$ are $N_E \times N_T$ matrices, and $\boldsymbol{X}$ is $N_S \times N_T$ matrix.

3.2. Mathematical formulation of well posed problems

The mathematical term "well-posed problem" stems from a definition given by Jacques Hadamard. It stats that the problem is well-posed if:

- a solution exists,
- the solution is unique,
- the solution's behavior changes continuously with the initial conditions. More precisely, small changes in measurements lead to small changes in solution.

The M/EEG inverse problem is to find $\boldsymbol{x}$ from Eq. (3.1) (or from Eq. (3.2)) with given $\boldsymbol{y}$ and $\boldsymbol{L}$. Let us analyse the well-posedness of this problem, first in a noise-free case, i. e. $\boldsymbol{b} \equiv 0$. We will also assume that $\boldsymbol{L}$ is a full-rank matrix.

1. *Existence of solution.*

   We can guarantee the existence of a solution if matrix $\boldsymbol{L}$ is square of rank N_E or has more columns than rows. It means that if there are more equations than variables, in general, there is no solution to the system. We can give a geometrical interpretation of this case. Each line of the system represents a hyper-plane in a space, which dimension is the number of variables (columns of $\boldsymbol{L}$). If the number of hyper-planes is higher then the dimension of the space, they do not have, in general, a common intersection point, i. e. the solution of the system does not exist (Fig. 3.1a).

2. *Uniqueness of solution.*

   The solution is unique if and only if $\boldsymbol{L}$ is a square and a full-rank matrix. If it has more columns than rows, there is an infinite number of solutions. More precisely, all elements from the null space of $\boldsymbol{L}$ can be added to a solution of the system to form a new one. A geometrical interpretation is that there are less hyper-planes than the dimension of the space and so their intersection is not a single point but a affine subspace (Fig. 3.1b).

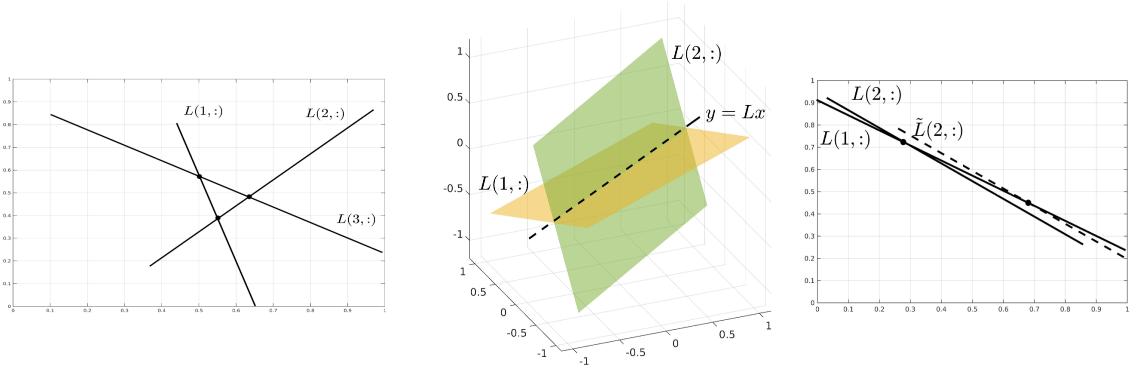
3. *Stability.*

   Even if $\boldsymbol{L}$ is square and invertible, the inversion can be unstable, i. e. small changes in measurements lead to large changes in source estimation. More precisely, let assume small changes in measurements $\delta\boldsymbol{y}$ and corresponding changes in sources $\delta\boldsymbol{x}$. From $\boldsymbol{y} = \boldsymbol{L}\boldsymbol{x}$, we have $\|\boldsymbol{y}\| \leq \|\boldsymbol{L}\|_2\|\boldsymbol{x}\|$, and so $\frac{1}{\|\boldsymbol{x}\|} \leq \frac{\|\boldsymbol{L}\|_2}{\|\boldsymbol{y}\|}$, where $\|\boldsymbol{L}\|_2$ denotes the 2-norm of operator $\boldsymbol{L}$ and is equal to the maximum singular value of $\boldsymbol{L}$: $\|\boldsymbol{L}\|_2 = \sigma_{max}(\boldsymbol{L})$. From $\delta\boldsymbol{y} = \boldsymbol{L}\delta\boldsymbol{x}$, we have $\|\delta\boldsymbol{x}\| \leq \|\boldsymbol{L}^{-1}\|_2\|\delta\boldsymbol{y}\|$. With $\|\boldsymbol{L}^{-1}\|_2 = \frac{1}{\sigma_{min}(\boldsymbol{L})}$, we get the inequality:

   $$\frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq \|\boldsymbol{L}\|_2\|\boldsymbol{L}^{-1}\|_2\frac{\|\delta\boldsymbol{y}\|}{\|\boldsymbol{y}\|} = \frac{\sigma_{max}(\boldsymbol{L})}{\sigma_{min}(\boldsymbol{L})}\frac{\|\delta\boldsymbol{y}\|}{\|\boldsymbol{y}\|}.$$

   The value $k(\boldsymbol{L}) = \frac{\sigma_{max}(\boldsymbol{L})}{\sigma_{min}(\boldsymbol{L})}$ is called **the condition number** of matrix $\boldsymbol{L}$. Matrices with high condition number are called ill-conditioned, meaning that the inversion of such a matrix is unstable. See Fig. 3.1c for a geometrical interpretation of an ill-conditioned system.

Note: with the presence of noise, i.e. $\boldsymbol{b} \not\equiv 0$, the problem automatically becomes ill-posed as it has more unknowns than equations whatever the size of $\boldsymbol{L}$.



(a) **(b)** **(c)**

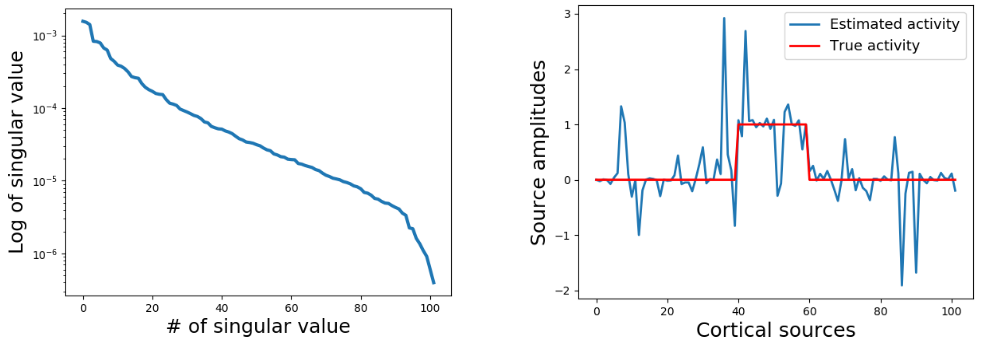
Figure 3.1: (a) $\boldsymbol{L}$ is a 3×2 matrix. 3 hyper-planes, in general, do not intersect at a point in a two dimensional space, so the inverse problem does not have a solution. (b) $\boldsymbol{L}$ is a 2×3 matrix. The intersection of two hyper-planes in three-dimensional space is a one-dimensional affine subspace (dashed line on the graph), not a single point. Thus the inverse problem has an infinite number of solutions. (c) $\boldsymbol{L}$ is a 2×2 invertible matrix, whose rows are so correlated that slightly moving the hyper-planes results in large change of the intersection point. The matrix $\boldsymbol{L}$ is ill-conditioned.

3.3. Ill-posedness of M/EEG inverse problem

The M/EEG inverse problem is ill-posed because of several reasons. First of all it is a fact that a lead field matrix $\boldsymbol{L}$, usually, has more columns (number of potential sources, which can exceed several thousand) than rows (number of sensors, rarely exceeding a few hundred). Moreover the lead field matrix is ill-conditioned and there is also the presence of noise in the measurements.

To demonstrate how ill-posed the M/EEG inverse problem is, we computed the MEG lead field for a sample subject from MNE-Python using their BEM solver with brain, scull and scalp surfaces computed from the subject's MRI. 102 magnetometers and more than 20000 cortical sources result in a very under-determined system which has an infinite number of solutions whatever the measurements are. But even if we randomly select 102 columns of the lead field, the resulting square matrix has a large condition number and its inversion is extremely unstable. Fig. 3.2a shows the singular values of such a matrix in a logarithmic scale. The singular values decrease exponentially fast which results in a large condition number. Fig. 3.2b shows the effect of inverting such a matrix - a small changes in sensor data result in a large differences in source reconstruction.

So in order to compute meaningful solutions, the M/EEG inverse problem needs to be approximated by a well-posed problem, which is usually done by introducing additional constraints to the problem.



(a) (b)

Figure 3.2: (a) Singular values of a randomly selected 102×102 submatrix of the lead field. (b) Red graph represent the simulated source amplitude. Sensor data were generated by applying 102×102 lead field matrix. Small noise was added to the sensor data (the norm of the noise was set to 10^{-5} of the signal norm). The blue graph represents the source amplitudes estimated by applying the inverse of the lead field matrix. As we can see, small changes in sensor data result in large differences in the source reconstruction.

3.4. Dipole fitting approach

Dipole fitting approaches assume that the measurement representing local cortical activity can be can be explained by a single dipole source. The objective is to find the location, orientation and amplitude of this dipole. Let us denote by $\boldsymbol{r}$ the dipole location, $\boldsymbol{q}$ its orientation and x its amplitude. The dipole fitting approach consists in minimizing the following data fitting error:

$$\min_{\boldsymbol{r},\boldsymbol{q},x} \|\boldsymbol{y} - \boldsymbol{L}(\boldsymbol{r},\boldsymbol{q})x\|_2^2,$$

where $\boldsymbol{L}(\boldsymbol{r},\boldsymbol{q})$ is a lead field generated by the dipole with position $\boldsymbol{r}$ and orientation $\boldsymbol{q}$ with unit amplitude. Noticing that $\boldsymbol{L}(\boldsymbol{r},\boldsymbol{q})$ is linear with respect to $\boldsymbol{q}$, the problem can be rewritten as:

$$\min_{\boldsymbol{r},\boldsymbol{x}} \|\boldsymbol{y} - \sum_{i=1}^{3} \boldsymbol{L}_i(\boldsymbol{r})x_i\|_2^2,$$

where $\boldsymbol{L}_i$ represent leadfields associated to the unit basis, and x_i are their corresponding amplitudes. For a fixed position, it is a least-square problem, which can be solved using the pseudo-inverse of $[\boldsymbol{L}_1, \boldsymbol{L}_2, \boldsymbol{L}_3]$.

This approach might be generalized to fit the data with several active dipoles. In this case, all possible dipole combinations are tested to fit the data. The problem of this approach is its exponential complexity and thus only a small number of dipoles can be assumed.

3.5. The variational approach to regularization methods

The variational approach to regularization methods became very popular in the last decade. It allows for an intuitive approach to modeling, gives a framework for its basic analysis, and can be solved using a variety of computational methods, in particular in the case of convex regularization functionals. The key idea to construct a variational regularization method for Eq. (3.1) consists in finding two functionals: a data fidelity term $F(\boldsymbol{x})$ measuring the distance between $\boldsymbol{y}$ and $\boldsymbol{L}\boldsymbol{x}$, and a regularization functional $R(\boldsymbol{x})$ favouring appropriate minimizers penalizing potential solutions with undesired structures [Benning and Burger, 2018]. The regularized solution is thus defined as :

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmin}}\, F(\boldsymbol{x}) + R(\boldsymbol{x}).$$

We will see in the next section that the choice of $F(\boldsymbol{x})$ depends on the noise properties. In the context of M/EEG, Euclidean norm is a reasonable choice for this metric. It is due to the fact that the Gaussion noise model is used for M/EEG measurments (see Section 3.6 for more details). A regularization parameter λ is often considered in the model to control the influence of the two terms in the minimizer. We obtain the following regularization problem:

$$\hat{\boldsymbol{x}}_\lambda = \underset{\boldsymbol{x}}{\operatorname{argmin}}\, \|\boldsymbol{y} - \boldsymbol{L}\boldsymbol{x}\|_2^2 + \lambda R(\boldsymbol{x}), \tag{3.3}$$

where:

- $\|\boldsymbol{y} - \boldsymbol{L}\boldsymbol{x}\|_2^2$ is a data fidelity term. This assumes that the noise $\boldsymbol{b}$ is Gaussian white. If, however, the noise is correlated, whitening of the data and of the lead field should be performed before solving the inverse problem [Engemann and Gramfort, 2015].

- $R(\boldsymbol{x})$ is a penalization or regularization term which encodes some a priori constraints on $\boldsymbol{x}$,

- λ is a regularization hyperparameter, which controls the relative impact of each term into the optimization. The higher λ, the more the solution is regularized.

There exists a wide range of methods which propose different choices of regularization term $R(\boldsymbol{x})$ to encode different a priori hypotheses about the source amplitudes $\boldsymbol{x}$.

3.5.1. l_2-norm regularization

The most common approach is based on the regularization of the Euclidean norm, i.e. $R(\boldsymbol{x}) = \|\boldsymbol{x}\|_2^2$. Called *Tikhonov regularization*, it is also known as *ridge regression* [Hastie et al., 2009] in statistics and *minimum norm estimates (MNE)* [Hämäläinen and Ilmoniemi, 1994] in the M/EEG community. The optimization problem becomes:

$$\hat{\boldsymbol{x}}_\lambda = \underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{L}\boldsymbol{x}\|_2^2 + \lambda\|\boldsymbol{x}\|_2^2. \tag{3.4}$$

This problem has a closed form unique solution:

$$\hat{\boldsymbol{x}}_\lambda = \boldsymbol{M}\boldsymbol{y} = (\boldsymbol{L}^T\boldsymbol{L} + \lambda\boldsymbol{I})^{-1}\boldsymbol{L}^T\boldsymbol{y} = \boldsymbol{L}^T(\boldsymbol{L}\boldsymbol{L}^T + \lambda\boldsymbol{I})^{-1}\boldsymbol{y},$$

where $\boldsymbol{I}$ is an identity matrix. Let us notice that it is better to use the formula which involves the inversion of matrix $(\boldsymbol{L}\boldsymbol{L}^T + \lambda\boldsymbol{I})$ because its size is the squared number of sensors, compared to the size of matrix $(\boldsymbol{L}\boldsymbol{L}^T + \lambda\boldsymbol{I})$ which is the squared number of sources.

Let $\boldsymbol{L} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T$ be the singular value decomposition (SVD) of $\boldsymbol{L}$. Then $\boldsymbol{L}\boldsymbol{L}^T = \boldsymbol{U}\boldsymbol{S}^2\boldsymbol{U}^T$. We may also present identity matrix as $\boldsymbol{I} = \boldsymbol{U}\boldsymbol{U}^T$. This leads to:

$$\hat{\boldsymbol{x}}_\lambda = \boldsymbol{L}^T\left(\boldsymbol{U}(\boldsymbol{S}^2 + \lambda\boldsymbol{I})\boldsymbol{U}^T\right)^{-1}\boldsymbol{y} = \boldsymbol{L}^T\boldsymbol{U}\boldsymbol{S}_\lambda^{-1}\boldsymbol{U}^T\boldsymbol{y},$$

where $\boldsymbol{S}_\lambda^{-1}$ is a diagonal matrix whose i-th diagonal element equals $\dfrac{1}{\sigma_i^2 + \lambda}$, $\{\sigma_i\}_i$ denoting the singular values of $\boldsymbol{L}$. We get:

$$\hat{\boldsymbol{x}}_\lambda = \sum_i \frac{\sigma_i}{\sigma_i^2 + \lambda}[U^T y]_i \boldsymbol{V}_i,$$

where $\boldsymbol{V}_i$ denotes the i-th column of matrix $\boldsymbol{V}$, and $[U^T y]_i$ is the i-th element of vector $\boldsymbol{U}^T\boldsymbol{y}$. Thus, we can see that the solution of Eq. (3.4) can be efficiently computed as a weighted sum of the singular vectors $\boldsymbol{V}_i$.

There is a wide range of generalizations of MNE with the form:

$$\hat{\boldsymbol{x}}_\lambda = \underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{L}\boldsymbol{x}\|_2^2 + \lambda\|\boldsymbol{W}\boldsymbol{x}\|_2^2, \tag{3.5}$$

where a linear transform $\boldsymbol{W}\boldsymbol{x}$ is regularized instead of the source amplitudes. $\boldsymbol{W}$ may stand for different purposes. In Lin et al. [2006], it is a depth weighting matrix, which reduces the bias of MNE towards superficial sources. In the case of Low Resolution brain Electromagnetic Tomography (LORETA), $\boldsymbol{W}$ is a discrete spacial Laplaciean operator that achieves smoothness between neighboring sources.

If the source covariance matrix $\boldsymbol{R}$ is known or estimated, methods such as dynamic statistical parametric mapping (dSPM) [Dale et al., 2000] and standardized LORETA (sLORETA) [Pascual-Marqui, 2002] use $\boldsymbol{W} = \boldsymbol{R}^{-\frac{1}{2}}$.

In all these generalizations, the source estimate is a linear transform of the measurements. We can thus see how source estimates are related to the real sources in the noise-free case.

$$\hat{\boldsymbol{x}} = \boldsymbol{M}\boldsymbol{y} = \boldsymbol{M}\boldsymbol{L}\boldsymbol{x} = \boldsymbol{R}\boldsymbol{x},$$

where $\boldsymbol{R}$ is so-called *resolution matrix.*, which can be used to quantify the relationship between true and estimated sources. The diagonal elements of $\boldsymbol{R}$ indicate the sensitivity of each estimated source to itself, and off-diagonal elements quantify the degree to which estimated sources are affected by the signal from all other sources in the brain. In general, the M/EEG resolution matrix has non-zero off-diagonal elements, since it cannot have a rank greater than the number of sensors. These off-diagonal elements introduce the leakage or cross-talk in the EEG/MEG inverse solutions [Farahibozorg et al., 2018, Grave de Peralta Menendez et al., 1997]. Because of this, such methods are not well adapted for localization of focal sources, because the spatial leakage results in a large amount of false positive sources (Fig. 3.3). One way to get rid of them is to use an a posteriori thresholding [Maksymenko et al., 2017]. But there are methods for M/EEG inverse problem which are non-linear and better adapted for sparse or focal source configurations as they produce no leakage effect. Most of them are based on the l_1-norm regularization.

3.5.2. l_1-norm regularization

In this subsection, we will introduce an approach based on the l_1-norm regularization, i.e. $R(\boldsymbol{x}) = \|\boldsymbol{x}\|_1 = \sum_i |x_i|$. It is known as *Lasso regression* [Tibshirani, 1996, Hastie et al., 2009] in statistics and *minimum current estimates (MCE)* [Uutela et al., 1999] in the M/EEG community. The optimization problem becomes:

$$\hat{\boldsymbol{x}}_\lambda = \underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{L}\boldsymbol{x}\|_2^2 + \lambda\|\boldsymbol{x}\|_1. \tag{3.6}$$

The main motivation of using l_1-norm regularization is its relation with *sparsity.*

If we are interested in the most sparse solution of the system, it would be reasonable to regularize with the l_0 pseudo norm, which simply counts the number of non-zero elements:

$$\hat{\boldsymbol{x}}_\lambda = \underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{L}\boldsymbol{x}\|_2^2 + \lambda\|\boldsymbol{x}\|_0. \tag{3.7}$$

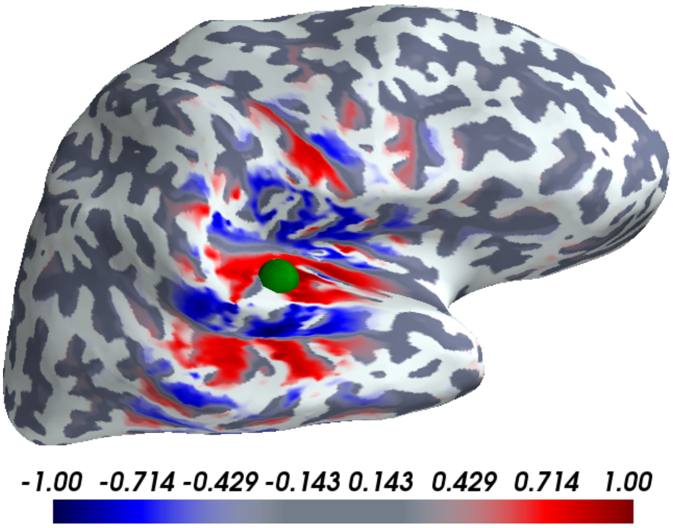


Figure 3.3: Simulated single dipole (green ball) and associated source estimates computed with sLORETA (color map). Source estimates are normalized so that maximum absolute value is 1, and threshold equal to 0.3 is used to get rid of small source values.

However the penalty function $\|\boldsymbol{x}\|_0$ is non-convex, and one can show that for a general matrix $\boldsymbol{L}$, even finding a solution that approximates the true minimum of Eq. (3.7) is NP-hard [Muthukrishnan, 2005].

The l_1-norm is a convex approximation of the l_0 pseudo norm [Davenport et al., 2012]. Problem (3.6) is convex and thus is computationally feasible. Even though l_1 regularization promotes sparsity, problems (3.6) and (3.7) are equivalent only under certain conditions on matrix $\boldsymbol{L}$, so-called the restricted isometry property (RIP) [Candes and Tao, 2005, J. Candès et al., 2006]. M/EEG lead field matrices are too ill-conditioned to verify this property. Thus solving Eq. (3.6) gives a sparse solution, but there is no guaranty that this solution is optimal in terms of Eq. (3.7).

The problem (3.6) is convex, but does not in general have a closed form solution. Several methods exist to find or approximate the solution of this problem. Least Angle Regression for Lasso (LARS) [Efron et al., 2004] is an algorithm which computes the entire solution path $\hat{\boldsymbol{x}}(\lambda)$ for all possible lambdas. It is possible because one can show that the solution of Eq. (3.6) is a piece-wise linear path with respect to λ and LARS is able to “follow” and explicitly compute the entire path.

It was shown that, in the worst case scenario, the number of linear segments (and thus the number of LARS iterations) is exponential in the number of variables [Mairal and Yu, 2012]. In practice, however, this method has near least squares complexity.

Another group of methods is based on approximating the solution of the problem (3.6). The most common approaches are proximal gradient descent (ISTA

[Daubechies et al., 2004] and FISTA [Beck and Teboulle, 2009]) and coordinate descent [Friedman et al., 2007]. The advantage of these methods is that, unlike LARS, they can be used for different Lasso generalizations and a large number of variables.

As in the case of MNE, the natural way to generalize l_1 regularization is to use it in the form:

$$\hat{\boldsymbol{x}}_\lambda = \operatorname*{argmin}_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{L}\boldsymbol{x}\|_2^2 + \lambda \|\boldsymbol{W}\boldsymbol{x}\|_1, \tag{3.8}$$

where the linear operator $\boldsymbol{W}$ may stand for different purposes. The idea is to explore sparseness in a transformed domain. If $\boldsymbol{W}$ is the discrete gradient on the cortical surface mesh, total variation regularization is used [Caselles et al., 2010, Ding, 2009, Clerc and Keriven, 2005, Adde et al., 2003], which promotes a piecewise constant cortical activity. In [Liao et al., 2012], $\boldsymbol{W}$ represents the spatial wavelet basis of cortical surface. In [Candès et al., 2008], $\boldsymbol{W}$ is a diagonal matrix whose diagonal elements represent weights of each source.

3.5.3. Further generalizations of regularization

l_1-norm and l_2-norm regularization are the the simplest, basic and most well studied approaches which can also serve as building blocks for other methods. In this subsection, we will provide references to such methods for further reading.

In [Lim et al., 2017], authors apply group-lasso [Yuan and Lin, 2006] to the M/EEG inverse problem. Given a cortical parcellation, the penalty term promotes inter-region sparsity via the l_1 norm and intra-region smoothness via the l_2 norm. A similar idea of combined l_1 and l_2 norms is used in [Gramfort et al., 2012]. So-called mixed-norm estimates methods (MxNE) can promote spatially focal sources with smooth temporal estimates. Authors expand the mixed norm idea for also taking into account time-frequency decomposition of the signal in [Gramfort et al., 2013].

For better taking into account the time component of M/EEG data, several methods include a multivariate autoregressive (MAR) model into the regularization problem [Fukushima et al., 2015, Belaoucha and Papadopoulo, 2017].

In [Becker et al., 2017], authors combine l_1 and total variation regularization to promote a small number of cortical regions each of which has constant activity.

There are quite a few methods which try to overcome the limitations of Lasso for sparse signal recovery (i.e. the fact that Lasso, in general, does not provide the sparsest solution). Thus, reweighted l_1 minimization was proposed in [Candès et al., 2008], which in many situations outperforms standard l_1 minimization. It has also been proved that non-convex penalties like l_q -norm with $q < 1$ penalty are able to recover sparsity in a more efficient way than Lasso, and methods were proposed to find local minima of such non-convex problems [Gasso et al., 2009, Bekhti et al., 2018]. Another approach uses a non-convex penalty presented in [Selesnick and Bayram, 2014]. The main idea of this approach is to make the regularization term as non-convex as possible, while keeping the whole minimization problem convex, and thus guaranteeing a unique global minimum. Unfortunately, the degree of non-convexity in the penalty term depends on the condition number of the lead field matrix, and thus, in the context of M/EEG, this approach does not differ much from the standard l_1 penalty.

3.6. Bayesian approach to M/EEG inverse problem

In this section, we will interpret the model (3.1) in terms of random variables. Thus, sources activity $\boldsymbol{x}$ and noise $\boldsymbol{b}$ are considered as realizations of random vectors X and B respectively. The measurement vector $\boldsymbol{y}$, as a function of random vectors, is also a realization of some random vector Y. The Bayesian approach to the M/EEG inverse problem consists in providing a priori hypotheses on X and B and applying an a posteriori probability (MAP) estimate to find $\hat{\boldsymbol{x}}$ [Bassett and Deride, 2019].

Let $p(\cdot)$ denote a probability density function, and $p(\cdot|\cdot)$ denote a conditional probability density. Given a Bayesian model and an observation $\boldsymbol{y}$, we define a posterior distribution of $\boldsymbol{x}$ through Bayes' rule:

$$p(\boldsymbol{x}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x})}{p(\boldsymbol{y})}.$$

A maximum a posteriori estimator is given by:

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmax}}\, p(\boldsymbol{x}|\boldsymbol{y}) = \underset{\boldsymbol{x}}{\operatorname{argmax}}\, p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x}).$$

Let us notice that $p(\boldsymbol{y})$ which does not depend on $\boldsymbol{x}$ can thus be excluded from the optimization problem. $p(\boldsymbol{y}|\boldsymbol{x})$ is the likelihood function and $p(\boldsymbol{x})$ encodes the a priori hypotheses on source configuration $\boldsymbol{x}$. It is often useful to consider the negative logarithm of the a posteriori probability, which results in:

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmin}} -\Big(\log p(\boldsymbol{y}|\boldsymbol{x}) + \log p(\boldsymbol{x})\Big). \tag{3.9}$$

3.6.1. Bayesian interpretation of l_1 and l_2 norm regularization

We start by describing the noise priors, which are the same for both l_1 and l_2-norm regularisations. We consider the noise to follow a centered Gaussian distribution. Let $\boldsymbol{\Sigma}_b$ denote the noise covariance matrix. Thus $B \sim \mathcal{N}(0, \boldsymbol{\Sigma}_b)$, and in the simple case when $\boldsymbol{\Sigma}_b = \sigma_b^2 \boldsymbol{I}$, where $\boldsymbol{I}$ is identity matrix, $B_i \sim \mathcal{N}(0, \sigma_b^2)$, where B_i is the i-th element of B.

We can now derive the distribution of Y, for a fixed $\boldsymbol{x}$: $Y|_{X=x} \sim \mathcal{N}(\boldsymbol{L}\boldsymbol{x}, \boldsymbol{\Sigma}_b)$. Thus, in the case when $\boldsymbol{\Sigma}_b = \sigma_b^2 \boldsymbol{I}$, the likelihood function equals:

$$p(\boldsymbol{y}|\boldsymbol{x}) = \frac{1}{\sqrt{2\pi|\boldsymbol{\Sigma}_b|}} \exp\Big(-\frac{1}{2}\|\boldsymbol{\Sigma}_b^{-\frac{1}{2}}(\boldsymbol{y} - \boldsymbol{L}\boldsymbol{x})\|_2^2 \Big).$$

Hypothesis about the a priori distribution of X differs between l_1 and l_2 approaches.

In *MNE* (l_2) sources are assumed to follow a centered Gaussian distribution with covariance $\boldsymbol{\Sigma}_x$, i. e. $X \sim \mathcal{N}(0, \boldsymbol{\Sigma}_x)$. Thus Eq. (3.9) becomes:

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{\Sigma}_b^{-\frac{1}{2}}(\boldsymbol{y} - \boldsymbol{L}\boldsymbol{x})\|_2^2 + \|\boldsymbol{\Sigma}_x^{-\frac{1}{2}}\boldsymbol{x}\|_2^2 \tag{3.10}$$

Let us notice that Eq. (3.5) and Eq. (3.10) are equivalent when $\boldsymbol{\Sigma}_b = \boldsymbol{I}$ (which is the case after whitening) and $\sqrt{\lambda}\boldsymbol{W} = \boldsymbol{\Sigma}_x^{-\frac{1}{2}}$.

In the case of *MCE* (l_1), assuming that each source independently follows a Laplace distribution, we get :

$$p(\boldsymbol{x}) = \prod_{i=1}^{N_S} \frac{1}{2\sigma_x} \exp\Big(-\frac{|x_i|}{\sigma_x} \Big) = \Big(\frac{1}{2\sigma_x} \Big)^{N_S} \exp\Big(-\frac{\|\boldsymbol{x}\|_1}{\sigma_x} \Big),$$

where N_S is the dimension of $\boldsymbol{x}$. Thus Eq. (3.9) becomes:

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{\Sigma}_b^{-\frac{1}{2}}(\boldsymbol{y} - \boldsymbol{L}\boldsymbol{x})\|_2^2 + \frac{1}{\sigma_x}\|\boldsymbol{x}\|_1. \tag{3.11}$$

Again, we may see that Eq. (3.11) is a particular case of the more general Eq. (3.6) with whitened data and lead field matrix.

3.6.2. Further generalizations of Bayesian approach

One of the possible generalizations of the Bayesian approach is to include a spatio-temporal constraint. In this case, the usual probabilistic model is extended by adding a level which describes the temporal evolution of neuronal current sources, for example using time-domain General Linear Models (GLMs) [Trujillo-Barreto et al., 2008].

In [Trujillo-Barreto et al., 2004], the Bayesian formulation offers a comparison framework for a wide range of available inverse methods. This allows to address the problem of model uncertainty that arises when dealing with different solutions for a single data set.

In [Friston et al., 2008], authors developed the multiple sparse priors (MSP) approach, in which they segment the brain into different pre-defined regions and promote all the dipoles in each region to be active or inactive jointly. In contrast, in [Wipf et al., 2010], authors developed the Champagne algorithm to promote activity to be concentrated on a sparse set of dipoles. In [Lucka et al., 2012], authors studied a hierarchical Bayesian model (HBM) offering significant improvements over established methods such as MNE and sLoreta.

A new hierarchical Bayesian model is proposed in [Costa et al., 2017a], with a multivariate Bernoulli Laplacian structured sparsity prior for brain activity. This distribution approximates a mixed $l_{2,0}$ pseudo norm on regularization term in a Bayesian framework.

In [Costa et al., 2017b] authors include the skull conductivity to the model as an unknown and propose a Bayesian algorithm to estimate the skull conductivity jointly with the brain activity directly from the M/EEG measurements.

Another class of methods based on the Bayesian technique are the Maximum Entropy on the Mean (MEM) [Clarke and Janday, 1989, Amblard et al., 2004] approach and its derivatives: wavelet based MEM (wMEM) dedicated to perform source localization of oscillatory patterns in the time-frequency domain [Lina et al., 2014]; ridge MEM (rMEM) dedicated to localize cortical sources exhibiting synchrony [Zerouali et al., 2013]; Standard MEM with stable clustering source localization (cMEM) [Chowdhury et al., 2013]; Connectivity Informed MEM (CIMEM) [Deslauriers-Gauthier et al., 2019], which estimates the directed dynamical functional connectivity whose delays are supported by the diffusion MRI derived structural connectivity.

3.7. Scanning methods

The methods presented in this section are called scanning methods because they assume that the activated dipole(s) are located at predefined positions. For example, the considered locations may belong to a regular sampling of the cortex region. The dipole orientations are either assumed to be known or can be left unconstrained. Scanning methods attempt to decide among the possible dipole locations where it is the most appropriate to place the sources so as to best describe the measurements.

3.7.1. Beamforming

Let $\boldsymbol{L}_i$ denote i-th column of lead field matrix $\boldsymbol{L}$, so that $\boldsymbol{y} = \sum_i \boldsymbol{L}_i x_i + \boldsymbol{b}$. For simplicity we assume the dipole orientations to be known. The method, however, can easily be generalized for unknown orientations. Every source amplitude x_i is assumed to be a random variable with variance σ_i. The sources are assumed be *uncorrelated* with each other and with noise [Van Veen and Buckley, 1988, van Drongelen et al., 1996]. Thus, the measurement $\boldsymbol{y}$ is also a random vector whose covariance matrix is:

$$\boldsymbol{C}_y = \sum_i \sigma_i \boldsymbol{L}_i \boldsymbol{L}_i^T + \boldsymbol{C}_b,$$

where $\boldsymbol{C}_b$ is the noise covariance.

For the i-th source, the objective is to construct an operator (a spatial filter) $\boldsymbol{W}_i \in \mathbf{R}^{N_E}$, such that $\hat{\boldsymbol{x}}_i = \boldsymbol{W}_i^T \boldsymbol{y}$. To get an ideal spatial filter, i. e. such that $\hat{x}_i = \boldsymbol{W}_i^T \boldsymbol{L}\boldsymbol{x} = x_i$, it should satisfy:

$$\boldsymbol{W}_i^T \boldsymbol{L}_j = \begin{cases} 0, & \text{if } i \neq j, \\ 1, & \text{if } i = j. \end{cases} \tag{3.12}$$

In the absence of noise, this would lead to a perfect reconstruction of the source activity. But in the context of the EEG/MEG signals, when $M > N$, such an ideal filter is not realizable. The idea of linearly constrained minimum variance (LCMV) filtering (aka beamforming) is to find $\boldsymbol{W}_i$ which minimizes the variance of the estimate ($\boldsymbol{W}_i^T \boldsymbol{C}_y \boldsymbol{W}_i$) while satisfying the constraint $\boldsymbol{W}_i^T \boldsymbol{L}_i = 1$. In other words, we replace the system (3.12) by an optimization problem:

$$\min_{\boldsymbol{W}_i} \boldsymbol{W}_i^T \boldsymbol{C}_y \boldsymbol{W}_i,$$
$$\text{s. t. } \boldsymbol{W}_i^T \boldsymbol{L}_i = 1.$$

The solution of this optimization problem can be obtained using Lagrange multipliers. As a result, the optimal filter is:

$$\boldsymbol{W}_i = \left(\boldsymbol{L}_i^T \boldsymbol{C}_y^{-1} \boldsymbol{L}_i\right)^{-1} \boldsymbol{L}_i^T \boldsymbol{C}_y^{-1}. \tag{3.13}$$

And the source estimate is computed as:

$$\hat{\boldsymbol{x}} = \boldsymbol{W}^T \boldsymbol{y},$$

where $\boldsymbol{W}_i$ is an i-th column of matrix $\boldsymbol{W}$.

Knowing the Eq. (3.13) of the spatial filter W_i for the i-th source, we can also derive the equation for the variance of this source:

$$\hat{x}_i^2 = \boldsymbol{W}_i^T \boldsymbol{y} \boldsymbol{y}^T \boldsymbol{W}_i = \boldsymbol{L}_i^T \boldsymbol{C}_y^{-1} \boldsymbol{L}_i.$$

This value, however, is biased by the noise, which introduces a spatially non-uniform component to the source variance estimate. It is due to the fact that the actual noise contribution for a particular source location depends on the lead field operator $\boldsymbol{L}$. To compensate for this, a renormalization is introduced:

$$\hat{x}_i^2 = \frac{\boldsymbol{L}_i^T \boldsymbol{C}_y^{-1} \boldsymbol{L}_i}{\boldsymbol{L}_i^T \boldsymbol{C}_b^{-1} \boldsymbol{L}_i}.$$

This value represents the ratio of the total activity to the normalized noise activity. Large values indicate regions of strong neural activity.

Note: as this method requires knowladge of the noise and the measurements covariance matrices, in practice, they are estimated from the data. Thus, it is necessary to have data with a sufficient number of time samples.

3.7.2. Multiple signal classification (MUSIC)

MUSIC corresponds to a family of algorithms which are based on the separation of the data space span($\boldsymbol{Y}$) into two mutually orthogonal subspaces, the signal space and the noise space. MUSIC is closely related to beamformers. One difference is that, unlike beamformers, MUSIC does not provide the time-courses during the localization process, they need to be estimated separately. Unlike beamformers, the MUSIC procedure does not require the inversion of the data covariance matrix. MUSIC also better tolerates time-correlated sources. Sources do not need to be uncorrelated, but they must not however be perfectly synchronised [Mosher et al., 1992, Mäkelä et al., 2018].

As for beamforming, it is necessary to have data with a sufficient number of time samples, for applying MUSIC algorithm. We assume the forward model with time dimension Eq. (3.2). Let denote by p the number of unsynchronised cortical sources. In practice, the signal subspace is estimated from the eigenvalue decomposition of the matrix $\boldsymbol{C} = \boldsymbol{Y}\boldsymbol{Y}^T = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^T$ by selecting the $\tilde{p}$ eigenvectors of $\boldsymbol{C}$, associated to the $\tilde{p}$ biggest eigenvalues. $\tilde{p}$ is an estimation of p, usually chosen to be larger than p. Assuming, that eigenvalues in the matrix D are sorted in decreasing order, let $\boldsymbol{U}_s = \boldsymbol{U}(:, 1 : \tilde{p})$. This matrix can be interpreted as a space of topographies, which can explain the signal.

Let $\boldsymbol{P}_s = \boldsymbol{U}_s\boldsymbol{U}_s^T$ be the orthogonal projection onto $\mathrm{span}(\boldsymbol{U}_s)$. MUSIC is based on the following property of this signal space projection $\boldsymbol{P}_s$: for any topography $\boldsymbol{l}$ $\|\boldsymbol{P}_s\boldsymbol{l}\| \leq \|\boldsymbol{l}\|$, while for any $\boldsymbol{l}$ from $\mathrm{span}(\boldsymbol{U}_s)$ $\boldsymbol{P}_s\boldsymbol{l} = \boldsymbol{l}$. The MUSIC localizer is then defined as follows:

$$\mu(\boldsymbol{l}) = \frac{\|\boldsymbol{P}_s\boldsymbol{l}\|^2}{\|\boldsymbol{l}\|^2}.$$

If the signal subspace is correctly estimated from data, $\mu(\boldsymbol{l}) \approx 1$ if $\boldsymbol{l}$ is one of the true source topographies, and $\mu(\boldsymbol{l}) < 1$ otherwise. So, the source locations, corresponding to the $\tilde{p}$ largest local maxima of μ, can in principle be found by computing $\mu(\boldsymbol{l})$ for each candidate-source location in the scanning grid, i. e. columns of lead field $\boldsymbol{L}$.

The selection of sources is however not a trivial problem. It can happen that the $\tilde{p}$ dipoles extracted may all explain the same part of the signal. To avoid this kind of situations, a recursive version of MUSIC called RAP-MUSIC (Recursively APplied MUSIC) has been proposed [Mosher and Leahy, 1998]. In this method, only a single maximum is extracted, and then the signal contribution of this source is subtracted from the measurements. The RAP-MUSIC method is then re-applied to this new measurement set.

More formally, after source topographies $\boldsymbol{l}_1, \ldots, \boldsymbol{l}_{k-1}$ have been found, we form the out-projector $\boldsymbol{Q}_k = \boldsymbol{I} - \boldsymbol{B}_k\boldsymbol{B}_k^\dagger$, where $\boldsymbol{B}_k = [\boldsymbol{l}_1, \ldots, \boldsymbol{l}_{k-1}]$ contains the topographies of previously found sources. The transformed signal space basis and its singular value decomposition are $\boldsymbol{Q}_k\boldsymbol{U}_s = \boldsymbol{U}_k\boldsymbol{D}_k\boldsymbol{V}_k^T$. We recompute the projection operator as $\boldsymbol{P}_k = \boldsymbol{U}_k\boldsymbol{U}_k^T$ and reapply MUSIC using $\boldsymbol{P}_k$ instead of $\boldsymbol{P}_s$ and $\boldsymbol{Q}_k\boldsymbol{l}$ instead of $\boldsymbol{l}$. The RAP-MUSIC locolizer at k-th iteration is:

$$\mu_k(\boldsymbol{l}) = \frac{\|\boldsymbol{P}_k\boldsymbol{Q}_k\boldsymbol{l}\|^2}{\|\boldsymbol{Q}_k\boldsymbol{l}\|^2}.$$

Truncated RAP-MUSIC (TRAP-MUSIC) method was recently proposed in [Mäkelä et al., 2018]. It corrects a hidden deficiency of the conventional RAP-MUSIC algorithm, which prevents estimation of the true number of brain-signal sources accurately.

Other MUSIC based methods have recently been developed to serve in specific applications, e.g., for locating extended sources (ExSo-MUSIC [Birot et al., 2011]), or for locating synchronous activity, either by exploiting source clustering (POP-MUSIC [Hesheng Liu and Schimpf, 2006]), or the imaginary part of the cross-spectrum (Wedge-MUSIC [Ewald et al., 2014], SC-MUSIC [Shahbazi et al., 2015]).

3.8. Conclusions

In this chapter, we gave an overview of several classes of the state-of-art M/EEG source localization methods. Here are some important points to highlight:

- The M/EEG inverse problem is ill-posed and usually has an infinite number of solutions.
- M/EEG lead field matrices are very ill-conditioned.
- Additional constraints are needed to get a "reasonable" solution of the problem. Each M/EEG source localization methods is thus based on specific a priori hypotheses.
- The majority of these methods are built to reconstruct a single solution to a M/EEG inverse problem. We will discuss in the Chapter 5 and Chapter 6 that it is not always a desired property.

Bibliography

G. Adde, M. Clerc, R. Keriven, and J. Kybic. Anatomy-based regularization for the inverse meeg problem. In *4th International Symposium on Noninvasive Functional Source Imaging within the human brain and heart*, Chieti, Sep 2003.

C. Amblard, E. Lapalme, and J. . Lina. Biomagnetic source detection by maximum entropy and graphical models. *IEEE Transactions on Biomedical Engineering*, 51(3):427–442, March 2004. ISSN 0018-9294. doi: 10.1109/TBME.2003.820999.

Robert Bassett and Julio Deride. Maximum a posteriori estimators as a limit of bayes estimators. *Mathematical Programming*, 174(1):129–144, Mar 2019. ISSN 1436-4646. doi: 10.1007/s10107-018-1241-0. URL `https://doi.org/10.1007/s10107-018-1241-0`.

Amir Beck and Marc Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM journal on imaging sciences*, 2(1):183–202, 2009.

H Becker, L Albera, P Comon, J.-C. Nunes, R Gribonval, J Fleureau, P Guillotel, and I Merlet. SISSY: An efficient and automatic algorithm for the analysis of EEG sources based on structured sparsity. *NeuroImage*, 157:157–172, 2017. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2017.05.046. URL `http://www.sciencedirect.com/science/article/pii/S1053811917304469`.

Yousra Bekhti, Felix Lucka, Joseph Salmon, and Alexandre Gramfort. A hierarchical Bayesian perspective on majorization-minimization for non-convex sparse regression: application to M/{EEG} source imaging. *Inverse Problems*, 34(8):85010, jun 2018. doi: 10.1088/1361-6420/aac9b3. URL `https://doi.org/10.1088{%}2F1361-6420{%}2Faac9b3`.

B. Belaoucha and T. Papadopoulo. Large brain effective network from EEG/MEG data and dmr information. In *2017 International Workshop on Pattern Recognition in Neuroimaging (PRNI)*, pages 1–4, June 2017. doi: 10.1109/PRNI.2017.7981511.

Martin Benning and Martin Burger. Modern regularization methods for inverse problems. *Acta Numerica*, 27:1–111, 2018. doi: 10.1017/S0962492918000016.

Gwénaël Birot, Laurent Albera, Fabrice Wendling, and Isabelle Merlet. Localization of extended brain sources from EEG/MEG: The ExSo-MUSIC approach. *NeuroImage*, 56(1):102–113, 2011. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2011.01.054. URL `http://www.sciencedirect.com/science/article/pii/S1053811911000905`.

E. J. Candes and T. Tao. Decoding by linear programming. *IEEE Transactions on Information Theory*, 51(12):4203–4215, Dec 2005. ISSN 0018-9448. doi: 10.1109/TIT.2005.858979.

Emmanuel J. Candès, Michael B. Wakin, and Stephen P. Boyd. Enhancing sparsity by reweighted l1 minimization. *Journal of Fourier Analysis and Applications*, 14 (5):877–905, Dec 2008. ISSN 1531-5851. doi: 10.1007/s00041-008-9045-x. URL `https://doi.org/10.1007/s00041-008-9045-x`.

Vicent Caselles, Antonin Chambolle, Daniel Cremers, Matteo Novaga, and Thomas Pock. An introduction to total variation for image analysis. *Theoretical Foundations and Numerical Methods for Sparse Recovery, De Gruyter, Radon Series Comp. Appl. Math.*, 9:263–340, 2010.

Rasheda Arman Chowdhury, Jean Marc Lina, Eliane Kobayashi, and Christophe Grova. Meg source localization of spatially extended generators of epileptic activity: Comparing entropic and hierarchical bayesian approaches. *PLOS ONE*, 8(2):1–19, 02 2013. doi: 10.1371/journal.pone.0055969. URL https://doi.org/10.1371/journal.pone.0055969.

C J S Clarke and B S Janday. The solution of the biomagnetic inverse problem by maximum statistical entropy. *Inverse Problems*, 5(4):483–500, aug 1989. doi: 10.1088/0266-5611/5/4/005. URL https://doi.org/10.1088%2F0266-5611%2F5%2F4%2F005.

Maureen Clerc and Renaud Keriven. Imaging methods for meg/eeg inverse problem. *International Journal of Bioelectromagnetism*, 7, 01 2005.

F. Costa, H. Batatia, T. Oberlin, and J. Tourneret. Skull conductivity estimation for EEG source localization. *IEEE Signal Processing Letters*, 24(4):422–426, April 2017b. ISSN 1070-9908. doi: 10.1109/LSP.2017.2669101.

Facundo Costa, Hadj Batatia, Thomas Oberlin, Carlos D'Giano, and Jean-Yves Tourneret. Bayesian EEG source localization using a structured sparsity prior. *NeuroImage*, 144:142–152, 2017a. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2016.08.064. URL http://www.sciencedirect.com/science/article/pii/S1053811916304554.

Anders M Dale, Arthur K Liu, Bruce R Fischl, Randy L Buckner, John W Belliveau, Jeffrey D Lewine, and Eric Halgren. Dynamic Statistical Parametric Mapping: Combining fMRI and MEG for High-Resolution Imaging of Cortical Activity. *Neuron*, 26(1):55–67, 2000. ISSN 0896-6273. doi: https://doi.org/10.1016/S0896-6273(00)81138-1. URL http://www.sciencedirect.com/science/article/pii/S0896627300811381.

I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57(11):1413–1457, 2004. doi: 10.1002/cpa.20042. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.20042.

Mark A. Davenport, Marco F. Duarte, Yonina C. Eldar, and Gitta Kutyniok. *Introduction to compressed sensing*, page 1–64. Cambridge University Press, 2012. doi: 10.1017/CBO9780511794308.002.

Samuel Deslauriers-Gauthier, Jean-Marc Lina, Russell Butler, Kevin Whittingstall, Guillaume Gilbert, Pierre-Michel Bernier, Rachid Deriche, and Maxime Descoteaux. White matter information flow mapping from diffusion mri and eeg. *NeuroImage*, 201:116017, 2019. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2019.116017. URL http://www.sciencedirect.com/science/article/pii/S1053811919305981.

Lei Ding. Reconstructing cortical current density by exploring sparseness in the transform domain. *Physics in Medicine and Biology*, 2009. ISSN 00319155. doi: 10.1088/0031-9155/54/9/006.

Bradley Efron, Trevor Hastie, Iain Johnstone, and Robert Tibshirani. Least angle regression. *Ann. Statist.*, 32(2):407–499, 04 2004. doi: 10.1214/009053604000000067. URL `https://doi.org/10.1214/009053604000000067`.

Denis A Engemann and Alexandre Gramfort. Automated model selection in covariance estimation and spatial whitening of MEG and EEG signals. *NeuroImage*, 108:328–342, 2015. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2014.12.040. URL `http://www.sciencedirect.com/science/article/pii/S1053811914010325`.

Arne Ewald, Forooz Shahbazi Avarvand, and Guido Nolte. Wedge MUSIC: A novel approach to examine experimental differences of brain source connectivity patterns from EEG/MEG data. *NeuroImage*, 101:610–624, 2014. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2014.07.011. URL `http://www.sciencedirect.com/science/article/pii/S1053811914005837`.

Seyedeh-Rezvan Farahibozorg, Richard N Henson, and Olaf Hauk. Adaptive cortical parcellations for source reconstructed EEG/MEG connectomes. *NeuroImage*, 169:23–45, 2018. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2017.09.009. URL `http://www.sciencedirect.com/science/article/pii/S1053811917307474`.

Jerome Friedman, Trevor Hastie, Holger Höfling, and Robert Tibshirani. Pathwise coordinate optimization. *Ann. Appl. Stat.*, 1(2):302–332, 12 2007. doi: 10.1214/07-AOAS131. URL `https://doi.org/10.1214/07-AOAS131`.

Karl Friston, Lee Harrison, Jean Daunizeau, Stefan Kiebel, Christophe Phillips, Nelson Trujillo-Barreto, Richard Henson, Guillaume Flandin, and Jérémie Mattout. Multiple sparse priors for the M/EEG inverse problem. *NeuroImage*, 39(3):1104–1120, 2008. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2007.09.048. URL `http://www.sciencedirect.com/science/article/pii/S1053811907008786`.

Makoto Fukushima, Okito Yamashita, Thomas R. Knösche, and Masa aki Sato. Meg source reconstruction based on identification of directed source interactions on whole-brain anatomical networks. *NeuroImage*, 105:408 – 427, 2015. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2014.09.066. URL `http://www.sciencedirect.com/science/article/pii/S1053811914008088`.

G. Gasso, A. Rakotomamonjy, and S. Canu. Recovering sparse signals with a certain family of nonconvex penalties and dc programming. *IEEE Transactions on Signal Processing*, 57(12):4686–4698, Dec 2009. ISSN 1053-587X. doi: 10.1109/TSP.2009.2026004.

A. Gramfort, D. Strohmeier, J. Haueisen, M. S. Hämäläinen, and M. Kowalski. Time-frequency mixed-norm estimates: Sparse M/EEG imaging with non-stationary source activations. *NeuroImage*, 70:410–422, 2013. ISSN 10538119. doi: 10.1016/j.neuroimage.2012.12.051. URL `http://dx.doi.org/10.1016/j.neuroimage.2012.12.051`.

Alexandre Gramfort, Matthieu Kowalski, and Matti Hämäläinen. Mixed-norm estimates for the M/{EEG} inverse problem using accelerated gradient methods. *Physics in Medicine and Biology*, 57(7):1937–1961, mar 2012. doi: 10.1088/0031-9155/57/7/1937. URL `https://doi.org/10.1088{%}2F0031-9155{%}2F57{%}2F7{%}2F1937`.

Rolando Grave de Peralta Menendez, Olaf Hauk, Sara Gonzalez Andino, Hermann Vogt, and Christoph Michel. Linear inverse solutions with optimal resolution kernels applied to electromagnetic tomography. *Human Brain Mapping*, 5(6):454–467, 1997. doi: 10.1002/(SICI)1097-0193(1997)5:6<454::AID-HBM6>3.0.CO;2-2. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291097-0193%281997%295%3A6%3C454%3A%3AAID-HBM6%3E3.0.CO%3B2-2`.

M. S. Hämäläinen and R. J. Ilmoniemi. Interpreting magnetic fields of the brain: minimum norm estimates. *Medical & Biological Engineering & Computing*, 32 (1):35–42, 1994. ISSN 01400118. doi: 10.1007/BF02512476.

T. Hastie, R. Tibshirani, and J.H. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer series in statistics. Springer, 2009. ISBN 9780387848846. URL `https://books.google.fr/books?id=eBSgoAEACAAJ`.

Hesheng Liu and P. H. Schimpf. Efficient localization of synchronous eeg source activities using a modified rap-music algorithm. *IEEE Transactions on Biomedical Engineering*, 53(4):652–661, April 2006. ISSN 0018-9294. doi: 10.1109/TBME.2006.870236.

Emmanuel J. Candès, Justin K. Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59, 08 2006. doi: 10.1002/cpa.20124.

Ke Liao, Min Zhu, Lei Ding, Sébastien Valette, Wenbo Zhang, and Deanna Dickens. Sparse imaging of cortical electrical current densities via wavelet transforms. *Physics in Medicine and Biology*, 57(21):6881–6901, oct 2012. doi: 10.1088/0031-9155/57/21/6881. URL `https://doi.org/10.1088{%}2F0031-9155{%}2F57{%}2F21{%}2F6881`.

Michael Lim, Justin M. Ales, Benoit R. Cottereau, Trevor Hastie, and Anthony M. Norcia. Sparse eeg/meg source estimation via a group lasso. *PLOS ONE*, 12(6): 1–29, 06 2017. doi: 10.1371/journal.pone.0176835. URL `https://doi.org/10.1371/journal.pone.0176835`.

Fa-Hsuan Lin, Thomas Witzel, Seppo P Ahlfors, Steven M Stufflebeam, John W Belliveau, and Matti S Hämäläinen. Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. *NeuroImage*, 31(1):160–171, 2006. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2005.11.054. URL `http://www.sciencedirect.com/science/article/pii/S1053811905024973`.

J. M. Lina, R. Chowdhury, E. Lemay, E. Kobayashi, and C. Grova. Wavelet-based localization of oscillatory sources from magnetoencephalography data. *IEEE Transactions on Biomedical Engineering*, 61(8):2350–2364, Aug 2014. ISSN 0018-9294. doi: 10.1109/TBME.2012.2189883.

Felix Lucka, Sampsa Pursiainen, Martin Burger, and Carsten H Wolters. Hierarchical Bayesian inference for the EEG inverse problem using realistic FE head models: Depth localization and source separation for focal primary currents. *NeuroImage*, 61(4):1364–1382, 2012. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2012.04.017. URL `http://www.sciencedirect.com/science/article/pii/S1053811912003989`.

Julien Mairal and Bin Yu. Complexity analysis of the lasso regularization path. *arXiv preprint arXiv:1205.0079*, 2012.

Niko Mäkelä, Matti Stenroos, Jukka Sarvas, and Risto J. Ilmoniemi. Truncated RAP-MUSIC (TRAP-MUSIC) for MEG and EEG source localization. *NeuroImage*, 2018. ISSN 10959572. doi: 10.1016/j.neuroimage.2017.11.013.

Kostiantyn Maksymenko, Bernard Giusiano, Nicolas Roehri, Christian-G. Bénar, and Jean-Michel Badier. Strategies for statistical thresholding of source localization maps in magnetoencephalography and estimating source extent. *Journal of Neuroscience Methods*, 290:95–104, 2017. ISSN 0165-0270. doi: https://doi.org/10.1016/j.jneumeth.2017.07.015. URL `http://www.sciencedirect.com/science/article/pii/S0165027017302583`.

J. C. Mosher, P. S. Lewis, and R. M. Leahy. Multiple dipole modeling and localization from spatio-temporal meg data. *IEEE Transactions on Biomedical Engineering*, 39(6):541–557, June 1992. ISSN 0018-9294. doi: 10.1109/10.141192.

John C. Mosher and Richard M. Leahy. Recursive MUSIC: A framework for EEG and MEG source localization. *IEEE Transactions on Biomedical Engineering*, 1998. ISSN 00189294. doi: 10.1109/10.725331.

S. Muthukrishnan. *Data Streams: Algorithms and Applications*. Foundations and trends in theoretical computer science. Now Publishers, 2005. ISBN 9781933019147. URL https://books.google.fr/books?id=415loiMd_c0C.

Roberto D. Pascual-Marqui. Standardized low-resolution brain electromagnetic tomography (sloreta): technical details. *Methods and findings in experimental and clinical pharmacology*, 24 Suppl D:5–12, 2002.

I. W. Selesnick and İ. Bayram. Sparse signal estimation by maximally sparse convex optimization. *IEEE Transactions on Signal Processing*, 62(5):1078–1092, March 2014. ISSN 1053-587X. doi: 10.1109/TSP.2014.2298839.

Forooz Shahbazi, Arne Ewald, and Guido Nolte. Self-consistent music: An approach to the localization of true brain interactions from eeg/meg data. *NeuroImage*, 112:299 – 309, 2015. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2015.02.054. URL http://www.sciencedirect.com/science/article/pii/S105381191500155X.

Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996. doi: 10.1111/j.2517-6161.1996.tb02080.x. URL https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1996.tb02080.x.

Nelson J Trujillo-Barreto, Eduardo Aubert-Vázquez, and Pedro A Valdés-Sosa. Bayesian model averaging in EEG/MEG imaging. *NeuroImage*, 21(4):1300–1319, 2004. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2003.11.008. URL http://www.sciencedirect.com/science/article/pii/S1053811903007286.

Nelson J Trujillo-Barreto, Eduardo Aubert-Vázquez, and William D Penny. Bayesian M/EEG source reconstruction with spatio-temporal priors. *NeuroImage*, 39(1):318–335, 2008. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2007.07.062. URL http://www.sciencedirect.com/science/article/pii/S1053811907006921.

K. Uutela, M. Hämäläinen, and E. Somersalo. Visualization of magnetoencephalographic data using minimum current estimates. *NeuroImage*, 10(2):173–180, 1999. ISSN 10538119. doi: 10.1006/nimg.1999.0454.

W. van Drongelen, M. Yuchtman, B. D. Van Veen, and A. C. van Huffelen. A spatial filtering technique to detect and localize multiple sources in the brain. *Brain Topography*, 9(1):39–49, Sep 1996. ISSN 1573-6792. doi: 10.1007/BF01191641. URL `https://doi.org/10.1007/BF01191641`.

B. D. Van Veen and K. M. Buckley. Beamforming: a versatile approach to spatial filtering. *IEEE ASSP Magazine*, 5(2):4–24, April 1988. ISSN 0740-7467. doi: 10.1109/53.665.

David P Wipf, Julia P Owen, Hagai T Attias, Kensuke Sekihara, and Srikantan S Nagarajan. Robust Bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG. *NeuroImage*, 49(1): 641–655, 2010. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2009.06.083. URL `http://www.sciencedirect.com/science/article/pii/S105381190900696X`.

Ming Yuan and Yi Lin. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society Series B*, 68:49–67, 02 2006. doi: 10.1111/j.1467-9868.2005.00532.x.

Y. Zerouali, C. L. Herry, B. Jemel, and J. Lina. Localization of synchronous cortical neural sources. *IEEE Transactions on Biomedical Engineering*, 60(3): 770–780, March 2013. ISSN 0018-9294. doi: 10.1109/TBME.2011.2176938.

Chapter 4

Fast Approximation of EEG Forward Problem and Application to Tissue Conductivity Estimation

Contents

4.1. Overview

As seen in the Chapter 2 and Chapter 3, a conductor model of the head (lead field) plays a central role in accurate source localization. This model is built from a specification of conductivity distribution of the modeled tissue compartments (scalp, skull, cerebrospinal fluid, brain gray and white matter, etc.), which in turn is linked to tissue geometry. Because it is impractical to directly measure head tissue conductivities in vivo for a specific subject, default values are often used. One problem is that the brain-to-skull conductivity ratio reported in the literature varies from 4 to 80 [Rush and Driscoll, 1968, Gonçalves et al., 2003, Hoekema et al., 2003]. As skull conductivity greatly influences the solution of the forward problem [Vallaghé and Clerc, 2008], localizing brain sources using uncertain conductivity values leads to important errors [Akalin Acar and Makeig, 2013, Wang and Ren, 2013, Van Uitert et al., 2004, Pohlmeier et al., 1997]. Taking account of the composite structure of the human skull could improve the accuracy of EEG source analysis [Dannhauer et al., 2011]: in such a case, conductivity of each skull tissue should be estimated independently. A possible solution is to estimate tissue conductivities and cortical activity simultaneously [Costa et al., 2017, Akalin Acar et al., 2016, Lew et al., 2009, Vallaghé et al., 2007]. Doing this requires to solve the EEG forward problem for possibly many different conductivity configurations. In Chapter 2, we defined a lead field as a linear operator relating brain electrical activity to potentials on EEG electrodes. Computing a lead field requires a matrix inversion which is computationally intensive for realistic head geometry represented with meshes with a large number of vertices. Thus, the required time for computing a large number of solutions for different conductivities quickly becomes impractical.

One way to deal with this problem is to approximate lead fields using a relatively small set of precomputed solutions. The reduced basis method approximates the solution of parametrized PDEs [Hesthaven et al., 2016] by reducing the number of degrees of freedom for Galerkin projection based on a set of already computed exact solutions.

The method presented in [Costa et al., 2017] is limited to one unknown conductivity: elements of the lead field matrix are approximated using polynomial interpolation based on a set of precomputed values. With this approach, the complexity of the approximation would increase fast with the number of unknown conductivities.

In this chapter, we propose a fast lead field approximation method which is robust to head model complexity and to the number of unknown conductivities. Our method is inspired by the reduced basis method, but adapted to the particular structure of the EEG forward problem and the nature of its conductivity parameter space. In section 4.2, we discuss the motivation for a fast lead field approximation method and present our method to do it. We evaluate the performance of our algorithm on simulated and real EEG data in section 4.3.

4.2. Fast lead field approximation method

Let us first remind the notations and equations that we have introduced in Chapter 2 and which will be used in this chapter.

The EEG Lead field matrix is:

$$\boldsymbol{L} = \boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1}\boldsymbol{D}_{\boldsymbol{\sigma}}. \tag{4.1}$$

We will also use decompositions of matrices $\boldsymbol{H}_{\boldsymbol{\sigma}}$ and $\boldsymbol{D}_{\boldsymbol{\sigma}}$:

$$\boldsymbol{H}_{\boldsymbol{\sigma}} = \sum_{i=1}^{N_H} \gamma_i(\boldsymbol{\sigma})\overline{\boldsymbol{H}}_i, \tag{4.2}$$

$$\boldsymbol{D}_{\boldsymbol{\sigma}} = \sum_{i=1}^{N_D} \lambda_i(\boldsymbol{\sigma})\overline{\boldsymbol{D}}_i \ , \tag{4.3}$$

4.2.1. Motivation

As already mentioned in introduction, in this work we are interested in computing the EEG forward problem for possibly many different conductivity configurations (with potential application to tissue conductivity estimation). The numerical complexity of lead field computation is essentially due to two operations:

- computation of matrices $\overline{\boldsymbol{H}}_i, i = 1 \dots N_H$ and $\overline{\boldsymbol{D}}_i, i = 1 \dots N_D$,
- inversion of the head matrix $\boldsymbol{H}_{\boldsymbol{\sigma}}$.

Matrices $\overline{\boldsymbol{H}}_i$ and $\overline{\boldsymbol{D}}_i$ do not depend on conductivity, therefore they need only to be computed once. The inverse of $\boldsymbol{H}_{\boldsymbol{\sigma}}$, however, depends on $\boldsymbol{\sigma}$ (see Eq. (4.2)) and, in general, has to be recomputed for each conductivity configuration of interest. For realistic head models implying a large number of unknowns, the size of matrix $\boldsymbol{H}_{\boldsymbol{\sigma}}$ is relatively large. As a result, in the case of a large set of conductivity configurations, the time required for inversion of head matrices and therefore computing lead fields becomes impractical. In this context, it can be worthwhile to compute not the ***exact*** lead fields for all needed conductivity configurations, but an ***approximation*** thereof.

In this chapter, we propose a lead field approximation method based on the following ideas. First of all, we assume that the manifold of parametrized solutions

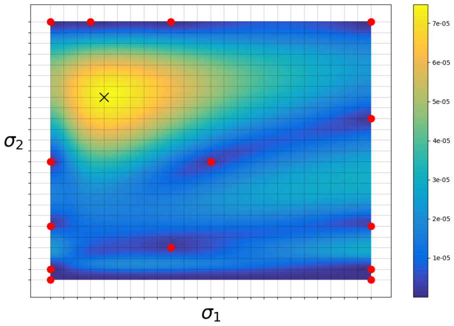


Figure 4.1: Schematic representation of a greedy support points selection. The domain of interest, here a two dimensional conductivity region, is sampled with a mesh grid. Red points correspond to current support points. The color represents upper bound error, computed at grid points (with sub-grid interpolation only for visual purposes). The grid point which maximises this error (black cross) is selected as the next support point. Let us notice that even if support points are selected from a finite sampling of the domain of interest, the lead field can be approximated for any conductivity configuration in the domain.

of EEG forward problem can be approximated by low-dimensional linear subspace. We choose a domain of interest in conductivity space, spanning a range of values for the head tissue conductivities. We then select ***support points*** within this domain, for which ***exact*** forward solutions will be computed. Based on these exact solutions a lead field ***approximation*** can be computed for any other conductivity configuration.

Support points are selected via a greedy algorithm. Similarly to the greedy basis generation used in reduced basis methods for parametrized partial differential equations [Hesthaven et al., 2016], our approach is an iterative procedure where each iteration adds one new support point. Each new support point is the one which maximizes an upper bound approximation error (to be defined), over a sampling of the domain of interest. Fig. 4.1 shows a schematic representation of this idea.

Remark: We make the assumption that the lead field as a function of conductivity is sufficiently regular so that the continuous conductivity domain can be explored using a sufficiently dense sampling.

In subsection 4.2.2, we formally define the approximation problem and the notion of upper bound for the approximation error. In subsection 4.2.3, we propose a particular choice for this error which is appropriate for the lead field approximation problem. Finally, in subsection 4.2.4, we describe the greedy algorithm for adding support points.

4.2.2. Approximation problem formulation

To approximate the true lead field matrix $\boldsymbol{L}(\boldsymbol{\sigma})$, we introduce coefficients $\boldsymbol{\alpha} \in \mathbb{R}^n$ where n is the number of support points. The lead field approximation is a matrix denoted by $\boldsymbol{L}_n(\boldsymbol{\alpha}, \boldsymbol{\sigma})$. We will see later that parameters $\boldsymbol{\alpha}$ represent coefficients of a linear approximation. We consider the following relative approximation error, using the Frobenius matrix norm $\|\boldsymbol{L}\|_F = (\sum l_{ij}^2)^{\frac{1}{2}}$:

$$E(\boldsymbol{L}(\boldsymbol{\sigma}), \boldsymbol{L}_n(\boldsymbol{\alpha}, \boldsymbol{\sigma})) = \frac{\|\boldsymbol{L}(\boldsymbol{\sigma}) - \boldsymbol{L}_n(\boldsymbol{\alpha}, \boldsymbol{\sigma})\|_F}{\|\boldsymbol{L}(\boldsymbol{\sigma})\|_F}. \tag{4.4}$$

To control this error without knowing $\boldsymbol{L}(\boldsymbol{\sigma})$, we introduce the definition of an upper bound for $E(\boldsymbol{L}, \boldsymbol{L}_n)$. The idea is to bound E with an error function E_n, which can be easily computed, converges to zero when the number of support points n increases, and thus ensures the convergence of E to zero.

Definition 1. *For $n \in \mathbb{N}$, function $E_n(\boldsymbol{\alpha}, \boldsymbol{\sigma})$ is an* ***upper bound approximation*** *of $E(\boldsymbol{L}(\boldsymbol{\sigma}), \boldsymbol{L}_n(\boldsymbol{\alpha}, \boldsymbol{\sigma}))$ if:*

$$\exists C > 0 \text{ such that } \forall \boldsymbol{\alpha} \in \mathbb{R}^n, \ E(\boldsymbol{L}(\boldsymbol{\sigma}), \boldsymbol{L}_n(\boldsymbol{\alpha}, \boldsymbol{\sigma})) \leq C \cdot E_n(\boldsymbol{\alpha}, \boldsymbol{\sigma}).$$

Let us remark that $C > 0$ is a constant which does not depend on $\boldsymbol{\alpha}$ but may depend on $\boldsymbol{\sigma}$.

As E is non-negative, it can be directly derived from definition 1 that if $E_n \xrightarrow[n\to\infty]{} 0$ then $E(\boldsymbol{L}, \boldsymbol{L}_n) \xrightarrow[n\to\infty]{} 0$.

Definition 2. *The sequence of matrices $\{\boldsymbol{L}_n(\boldsymbol{\alpha}_n^*(\boldsymbol{\sigma}), \boldsymbol{\sigma}), n \in \mathbb{N}\}$ is a* ***valid approximation*** *of $\boldsymbol{L}(\boldsymbol{\sigma})$ with* ***optimal approximation parameters*** *$\{\boldsymbol{\alpha}_n^*(\boldsymbol{\sigma}), n \in \mathbb{N}\}$ if both following conditions hold:*

(i) $\forall n, \ \boldsymbol{\alpha}_n^*(\boldsymbol{\sigma}) = \operatorname{argmin}_{\boldsymbol{\alpha}} E_n(\boldsymbol{\alpha}, \boldsymbol{\sigma})$,

(ii) $E_n(\boldsymbol{\alpha}_n^*(\boldsymbol{\sigma}), \boldsymbol{\sigma}) \xrightarrow[n\to\infty]{} 0.$

4.2.3. Choice of upper bound approximation

As mentioned previously, the objective of the approximation is to circumvent the inversion of matrix $\boldsymbol{H}_{\boldsymbol{\sigma}}$. One possible way to parametrize the approximation matrix $\boldsymbol{L}_n(\boldsymbol{\alpha}, \boldsymbol{\sigma})$ is to approximate the inverse $\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1}$ as a linear combination of precomputed inversions $\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1}$ at n support points:

$$\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1} \approx \sum_{i=1}^{n} \alpha_i \boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1} . \tag{4.5}$$

Based on Eq. (4.1), the proposed lead field approximation takes the form:

$$\boldsymbol{L}_n(\boldsymbol{\alpha}, \boldsymbol{\sigma}) = \sum_{i=1}^{n} \alpha_i \boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1} \boldsymbol{D}_{\boldsymbol{\sigma}} \ .$$

Using the linear decomposition of $\boldsymbol{D}_{\boldsymbol{\sigma}}$ in Eq. (4.3), we represent the lead field approximation as a linear combination of matrices which are independent of $\boldsymbol{\sigma}$:

$$\boldsymbol{L}_n(\boldsymbol{\alpha}, \boldsymbol{\sigma}) = \sum_{i=1}^{n} \sum_{j=1}^{N_D} \alpha_i \lambda_j(\boldsymbol{\sigma}) \boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1} \overline{\boldsymbol{D}}_j = \sum_{i=1}^{n} \sum_{j=1}^{N_D} \beta_{ij}(\boldsymbol{\sigma}, \boldsymbol{\alpha}) \overline{\boldsymbol{L}}_{ij} \ , \tag{4.6}$$

where $\overline{\boldsymbol{L}}_{ij} = \boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1} \overline{\boldsymbol{D}}_j$ are matrices which do not depend on the conductivity $\boldsymbol{\sigma}$ and thus can be precomputed, and $\boldsymbol{\beta}_{ij}(\boldsymbol{\sigma}, \boldsymbol{\alpha}) = \alpha_i \lambda_j(\boldsymbol{\sigma})$ are conductivity-dependent scalars.

As λ_j are known (see Section 2.6 in Chapter 2), the question is how to compute optimal coefficients $\boldsymbol{\alpha}$. We will base our approach on the property of an inverse matrix: $\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1} \cdot \boldsymbol{H}_{\boldsymbol{\sigma}} = \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix. In the context of the EEG forward problem, it is more relevant to consider matrix $\boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1}$, which has smaller dimensions than $\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1}$ and which contains only the relevant part of $\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1}$. Using $\boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1} \cdot \boldsymbol{H}_{\boldsymbol{\sigma}} = \boldsymbol{S}$ and Eq. (4.5), we could seek $\boldsymbol{\alpha}$ by minimizing the expression:

$$\boldsymbol{\alpha}_n^*(\boldsymbol{\sigma}) = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}}\, E_n(\boldsymbol{\alpha}, \boldsymbol{\sigma}) = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \left\| \boldsymbol{S} - \sum_{i=1}^{n} \alpha_i \boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1} \boldsymbol{H}_{\boldsymbol{\sigma}} \right\|_F \ . \tag{4.7}$$

Proposition 1 shows that this choice of $\boldsymbol{\alpha}$ leads to an upper bound approximation in the sense of Definition 1.

Proposition 1. *If matrices* $\{\boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1} \boldsymbol{H}_{\boldsymbol{\sigma}}\}_{i=1}^{n}$ *are linearly independent, then*

$$E_n(\boldsymbol{\alpha}, \boldsymbol{\sigma}) = \left\| \boldsymbol{S} - \sum_{i=1}^{n} \alpha_i \boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1} \boldsymbol{H}_{\boldsymbol{\sigma}} \right\|_F$$

is an upper bound approximation of error $E(\boldsymbol{L}, \boldsymbol{L}_n)$ *(see Eq. (4.4)) and* $\boldsymbol{\alpha}_n^*(\boldsymbol{\sigma})$ *is an optimal approximation parameter.*

Proof.

$$
\begin{aligned}
&\left\|\boldsymbol{L}(\boldsymbol{\sigma}) - \boldsymbol{L}_n(\boldsymbol{\alpha},\boldsymbol{\sigma})\right\|_F = \\
&\left\|\boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1}\boldsymbol{D}_{\boldsymbol{\sigma}} - \sum_i \alpha_i \boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1}\boldsymbol{D}_{\boldsymbol{\sigma}}\right\|_F = \\
&\left\|\left(\boldsymbol{S} - \sum_i \alpha_i \boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1}\boldsymbol{H}_{\boldsymbol{\sigma}}\right) H_{\boldsymbol{\sigma}}^{-1}\boldsymbol{D}_{\boldsymbol{\sigma}}\right\|_F \leq \\
&\left\|\boldsymbol{S} - \sum_i \alpha_i \boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1}\boldsymbol{H}_{\boldsymbol{\sigma}}\right\|_F \left\|\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1}\boldsymbol{D}_{\boldsymbol{\sigma}}\right\|_F = \\
&\left\|\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1}\boldsymbol{D}_{\boldsymbol{\sigma}}\right\|_F E_n(\boldsymbol{\alpha},\boldsymbol{\sigma}) \ .
\end{aligned}
$$

Now we can see that E_n is an upper bound of error from Eq. (4.4):

$$
E(\boldsymbol{L},\boldsymbol{L}_n) = \frac{\|\boldsymbol{L}-\boldsymbol{L}_n\|_F}{\|\boldsymbol{L}\|_F} \leq \frac{\|\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1}\boldsymbol{D}_{\boldsymbol{\sigma}}\|_F}{\|\boldsymbol{L}\|_F} E_n(\boldsymbol{\alpha},\boldsymbol{\sigma}) = C \cdot E_n(\boldsymbol{\alpha},\boldsymbol{\sigma}) \ .
$$

Let $\text{vec}(\boldsymbol{M})$ represent the vectorization of matrix $\boldsymbol{M}$. Using the fact that $\|\boldsymbol{M}\|_F = \|\text{vec}(\boldsymbol{M})\|_2$, problem (4.7) can be reformulated as:

$$
\begin{aligned}
\boldsymbol{\alpha}_{\boldsymbol{n}}^*(\boldsymbol{\sigma}) = \underset{\boldsymbol{\alpha}}{\text{argmin}} \left\|\boldsymbol{S} - \sum_i \alpha_i \boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1}\boldsymbol{H}_{\boldsymbol{\sigma}}\right\|_F &= \\
\underset{\boldsymbol{\alpha}}{\text{argmin}} \left\| v(\boldsymbol{S}) - \sum_i \alpha_i \,\text{vec}(\boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1}\boldsymbol{H}_{\boldsymbol{\sigma}})\right\|_2 &,
\end{aligned}
$$

which is a linear projection problem of $\text{vec}(\boldsymbol{S})$ onto the space spanned by the set $S_n = \{\text{vec}(\boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1}\boldsymbol{H}_{\boldsymbol{\sigma}})\}_{i=1}^n$. It means that if S_n is made of linearly independent vectors for $n_{max} = dim(\text{vec}(\boldsymbol{S}))$, then $E_{n_{max}}(\boldsymbol{\alpha}_{n_{max}}^*(\boldsymbol{\sigma}),\boldsymbol{\sigma}) = 0$. This directly implies that $E_n(\boldsymbol{\alpha}_n^*(\boldsymbol{\sigma}),\boldsymbol{\sigma}) \underset{n\to\infty}{\longrightarrow} 0$ so both conditions of Definition 2 are verified if S_n is a linearly independent set.

□

Linear dependence of ***basis matrices*** $\{\boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1}\boldsymbol{H}_{\boldsymbol{\sigma}}\}_{i=1}^n$, would mean that lead fields corresponding to some support points could be exactly represented as a linear combination of lead fields at other support points. So the linear independence condition of Proposition 1 can be handled by properly selecting support points. This will be discussed in detail in the next section.

We now express problem (4.7) as a simple least squares problem. Using the linear decomposition of $\boldsymbol{H}_{\boldsymbol{\sigma}}$ as a sum of conductivity-independent matrices (Eq. (4.2)), we get:

$$\boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1}\boldsymbol{H}_{\boldsymbol{\sigma}} = \sum_{j=1}^{N_H} \gamma_j(\boldsymbol{\sigma})\boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1}\overline{\boldsymbol{H}}_j \ .$$

We can then vectorize matrices $\boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1}\overline{\boldsymbol{H}}_j$ and assemble them as columns of a new matrix $\boldsymbol{K}$. Let $\boldsymbol{K}_{(i-1)N_H+j} = \mathrm{vec}(\boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1}\overline{\boldsymbol{H}}_j)$ denote the $((i-1)N_H+j)$-th column of $\boldsymbol{K}$. The matrix $\boldsymbol{K}$ will thus have $n \cdot N_H$ columns and $N_E \cdot N_V$ rows.

Let us also denote $\boldsymbol{\Gamma}_{\boldsymbol{\sigma}} = \boldsymbol{I} \otimes \boldsymbol{\gamma}(\boldsymbol{\sigma})$, where $\boldsymbol{\gamma}(\boldsymbol{\sigma}) = (\gamma_1(\boldsymbol{\sigma}), \cdots, \gamma_{N_H}(\boldsymbol{\sigma}))$ and $\boldsymbol{I}$ is an identity matrix of size n (number of support points):

$$\boldsymbol{\Gamma}_{\boldsymbol{\sigma}} = \begin{bmatrix} \gamma_1(\boldsymbol{\sigma}) & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{N_H}(\boldsymbol{\sigma}) & 0 & \dots & 0 \\ 0 & \gamma_1(\boldsymbol{\sigma}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \gamma_{N_H}(\boldsymbol{\sigma}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \gamma_1(\boldsymbol{\sigma}) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \gamma_{N_H}(\boldsymbol{\sigma}) \end{bmatrix}$$

Using the introduced notations, we can show that $\boldsymbol{K}\boldsymbol{\Gamma}_{\boldsymbol{\sigma}}$ represents a linear projection basis and that the solution $\boldsymbol{\alpha}_n^*(\boldsymbol{\sigma})$ of problem (4.7) is given by the solution of the system:

$$(\boldsymbol{\Gamma}_{\boldsymbol{\sigma}}^T\boldsymbol{K}^T\boldsymbol{K}\boldsymbol{\Gamma}_{\boldsymbol{\sigma}})\boldsymbol{\alpha}_n^*(\boldsymbol{\sigma}) = \boldsymbol{\Gamma}_{\boldsymbol{\sigma}}^T\boldsymbol{K}^T \,\mathrm{vec}(\boldsymbol{S}) \ . \tag{4.8}$$

Proof. Using the decomposition of matrix $\boldsymbol{H}_{\boldsymbol{\sigma}}$ (see Eq. (4.2)), we get:

$$\sum_{i=1}^{n} \alpha_i \boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1}\boldsymbol{H}_{\boldsymbol{\sigma}} = \sum_{i=1}^{n} \alpha_i \boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1}\Big(\sum_{j=1}^{N_H}\gamma_j(\boldsymbol{\sigma})\overline{\boldsymbol{H}}_j\Big) = \sum_{i=1}^{n}\alpha_i\sum_{j=1}^{N_H}\gamma_j(\boldsymbol{\sigma})\boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1}\overline{\boldsymbol{H}}_j \ .$$

Thus Eq. (4.7) becomes:

$$\begin{aligned}
&\boldsymbol{\alpha}_{\boldsymbol{n}}^{*}(\boldsymbol{\sigma}) = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \left\| \boldsymbol{S} - \sum_{i}^{n} \alpha_i \boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1}\boldsymbol{H}_{\boldsymbol{\sigma}} \right\|_F = \\
&\underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \left\| \operatorname{vec}(\boldsymbol{S}) - \sum_{i=1}^{n} \alpha_i \sum_{j=1}^{N_H} \gamma_j(\boldsymbol{\sigma}) \operatorname{vec}(\boldsymbol{S}\boldsymbol{H}_{\boldsymbol{\sigma}_i}^{-1}\overline{\boldsymbol{H}}_j) \right\|_2 = \\
&\underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \left\| \operatorname{vec}(\boldsymbol{S}) - \sum_{i=1}^{n} \alpha_i \sum_{j=1}^{N_H} \gamma_j(\boldsymbol{\sigma}) \boldsymbol{K}_{(i-1)N_H+j} \right\|_2 = \\
&\underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \left\| \operatorname{vec}(\boldsymbol{S}) - \boldsymbol{K}\boldsymbol{\Gamma}_{\boldsymbol{\sigma}}\ \boldsymbol{\alpha} \right\|_2 .
\end{aligned}$$

This is the projection problem of vec $\boldsymbol{S}$ onto the space spanned by the columns of $\boldsymbol{K}\boldsymbol{\Gamma}_{\boldsymbol{\sigma}}$. Its solution is given by Eq. (4.8).

□

Some important remarks must be noted:

1. The size of matrix $\boldsymbol{\Gamma}_{\boldsymbol{\sigma}}^T\boldsymbol{K}^T\boldsymbol{K}\boldsymbol{\Gamma}_{\boldsymbol{\sigma}}$ is $n \times n$. The complexity of problem (4.8) therefore does not depend on the size of the head model but only on the number of support points.

2. Matrices $\boldsymbol{K}^T\boldsymbol{K}$ and $\boldsymbol{K}^T \operatorname{vec}(\boldsymbol{S})$ are independent of $\boldsymbol{\sigma}$. Moreover, adding a new support point amounts to adding new elements to these matrices - they do not have to be fully recomputed.

To summarize, given a set of n support points, for each $\boldsymbol{\sigma}$ we use Eq. (4.8) to compute the optimal approximation parameter $\boldsymbol{\alpha}_n^*(\boldsymbol{\sigma})$ and then use Eq. (4.6) to compute lead field approximation $\boldsymbol{L}_n(\boldsymbol{\alpha}_n^*(\boldsymbol{\sigma}), \boldsymbol{\sigma})$.

4.2.4. Selection of support points

The method introduced in subsection 4.2.3 supposes that a set of support points is given allowing to precompute matrices $\boldsymbol{K}^T\boldsymbol{K}$, $\boldsymbol{K}^T \operatorname{vec}(\boldsymbol{S})$ and $\overline{\boldsymbol{L}}_{ij}$. This section answers the question of how to select them.

The idea is to start with a small number of support points and to add new points iteratively. At each iteration, we select a conductivity which maximizes the upper bound error on a discretization of the conductivity domain of interest. The procedure is summarized in Algorithm 1.

- In this work, the discretization of the domain of interest was uniform, but this choice is arbitrary: the sampling can be chosen according to the application.

Algorithm 1 Support points selection

```
Input: Sampling $\Sigma = \{\boldsymbol{\sigma}_i\}_{i=1}^{M}$ of the domain of interest ; set Supp of initial support points.
$n \leftarrow |\text{Supp}|$.
Compute matrices $\boldsymbol{K}^T\boldsymbol{K}$ and $\boldsymbol{K}^T \operatorname{vec}(\boldsymbol{S})$.
repeat
    for each $\boldsymbol{\sigma}$ in $\Sigma$ do
        Compute $\boldsymbol{\alpha}_n^*(\boldsymbol{\sigma})$ with Eq. (4.8).
        Compute $E_n(\boldsymbol{\alpha}_{\boldsymbol{n}}^*(\boldsymbol{\sigma}), \boldsymbol{\sigma})$.
    Supp = Supp $\cup \operatorname{argmax}_{\boldsymbol{\sigma}\in\Sigma}\{E_n(\boldsymbol{\alpha}_{\boldsymbol{n}}^*(\boldsymbol{\sigma}), \boldsymbol{\sigma})$
    Update matrices $\boldsymbol{K}^T\boldsymbol{K}$ and $\boldsymbol{K}^T \operatorname{vec}(\boldsymbol{S})$.
    $n \leftarrow n + 1$
until convergence
Output : Matrices $\boldsymbol{K}^T\boldsymbol{K}$ and $\boldsymbol{K}^T \operatorname{vec}(\boldsymbol{S})$.
```

- In this work we initialized the basis with the corners of the domain of interest (e.g. 4 support points for the $2D$ domain). This choice is arbitrary, and Algorithm 1 can be initialized with for instance a single support point. *Remark:* it is important to check that the initialized basis is a full rank matrix.

- There are several ways to define a convergence criterion. Even though the real approximation error is unknown, it seems reasonable to control the maximum value of the upper bound error, noted E_n^* (stop when $E_n^* < \epsilon$), or its decreasing rate (stop when $|E_n^* - E_{n-1}^*| < \epsilon$).

Algorithm 1 has interesting properties which make it more efficient than a random selection of support points.

First, the valid upper bound error is controlled, in a strong sense:

$$\forall \boldsymbol{\sigma} \;\; E_{n+1}(\boldsymbol{\alpha}_{n+1}^*(\boldsymbol{\sigma}), \boldsymbol{\sigma}) \leq E_n(\boldsymbol{\alpha}_n^*(\boldsymbol{\sigma}), \boldsymbol{\sigma}) \; .$$

This is true because the projection space at step n is a subspace of the one at step $n+1$. Also recall that $E_n(\boldsymbol{\alpha}_n^*(\boldsymbol{\sigma}), \boldsymbol{\sigma}) \underset{n\to\infty}{\longrightarrow} 0$, because we are dealing with finite spaces.

Second, Algorithm 1 ensures that Eq. (4.8) has a unique solution, i. e. that the matrix $\boldsymbol{K\Gamma_{\sigma}}$ has full rank. Let us analyze the two possible situations:

1. If $\forall n \leq M \; \max_{\boldsymbol{\sigma}\in\Sigma} E_n(\boldsymbol{\alpha}_{\boldsymbol{n}}^*(\boldsymbol{\sigma}), \boldsymbol{\sigma}) \neq 0$, each iteration adds a support point which has a strictly positive approximation error. This means that the matrix

$\boldsymbol{SH}_{\sigma^*}^{-1}$ associated to this point is linearly independent from the previous basis matrices $\{\boldsymbol{SH}_{\sigma_i}^{-1}\}_{i=1}^{n}$. Thus the updated projection basis $\boldsymbol{K\Gamma}_\sigma$ stays full rank.

2. Conversely for n such that $\max_{\boldsymbol{\sigma}\in\Sigma} E_n(\boldsymbol{\alpha}_{\boldsymbol{n}}^*(\boldsymbol{\sigma}),\boldsymbol{\sigma}) = 0$, then all lead fields from the discretized conductivity space can be ***exactly*** represented as a linear combination of the precomputed lead fields from support points. In practice, this naturally means that no additional support point is needed, unless new conductivity samples are introduced.

All these observations show that Algorithm 1 guarantees that the necessary condition for Proposition 1 is satisfied.

Remark: The fact that projection matrix $\boldsymbol{K\Gamma_\sigma}$ is full rank, and hence that matrix $\boldsymbol{\Gamma_\sigma^T K^T K \Gamma_\sigma}$ is invertible, ensures the continuity of $\boldsymbol{\alpha}_n^*(\boldsymbol{\sigma})$ with respect to conductivity. The upper bound approximation $E_n(\boldsymbol{\alpha}_{\boldsymbol{n}}^*(\boldsymbol{\sigma}),\boldsymbol{\sigma})$ is therefore also continuous in $\boldsymbol{\sigma}$ and hence it is bounded on any compact conductivity domain, where the exact solution exists.

Once support points have been selected, for any new conductivity $\boldsymbol{\sigma}$, the new lead field matrix is approximated with Algorithm 2.

Algorithm 2 Lead field approximation

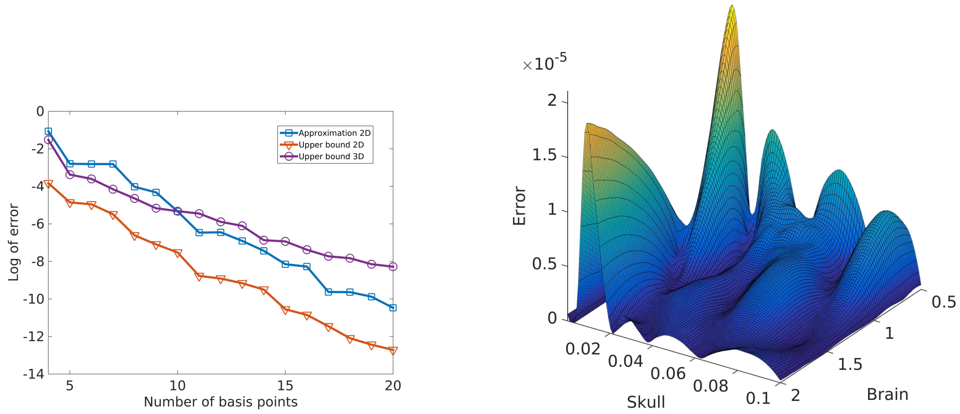
```
Input: $\boldsymbol{\sigma}$; matrices $\boldsymbol{K}^T\boldsymbol{K}$, $\boldsymbol{K}^T \operatorname{vec}(\boldsymbol{S})$ and $\{\overline{\boldsymbol{L}}_{ij}\}$, corresponding to support points set Supp.
    Compute $\boldsymbol{\alpha}_n^*(\boldsymbol{\sigma})$ with Eq. (4.8)
Output: $\boldsymbol{L}_n(\boldsymbol{\alpha}_n^*(\boldsymbol{\sigma}),\boldsymbol{\sigma})$ with Eq. (4.6).
```

4.3. Numerical results

4.3.1. Data description

To demonstrate numerically the proposed method, we use the data from [Tadel et al., 2016], which includes anatomical data and EEG data recorded with Yokogawa/KIT, and processed with Brainstorm [Tadel et al., 2011]. We use a realistic head model with three layers: brain, skull and scalp. Each surface is represented as a triangular mesh with 1082 vertices. The source space contains 15002 dipoles with fixed orientations (normal to the cortical surface).

We use the symmetric BEM implementation from OpenMEEG [Kybic et al., 2005, Gramfort et al., 2010]. The head matrix $\boldsymbol{H}$ has size 7566×7566. Matrix $\boldsymbol{D}$ has size 7566×15002.



(a) **(b)**

Figure 4.2: (a) Logarithm of maximal (over conductivity samples) upper bound and approximation errors for two and three unknown conductivities (conductivity space of dimension 2 and 3) as a function of number of support points; (b) Approximation error for 22 support points.

We use a model with 41 EEG electrodes, so the size of the selection matrix $\boldsymbol{S}$ is 41×7566 and the lead field is a 41×15002 matrix.

Conductivity of the scalp is taken equal to 1, while the brain and skull conductivities (σ_1 and σ_2 respectively) are variable and form the parameter space (a subspace of $\mathbb{R}^2$). We are interested in a subset of this space represented by a rectangle which spans values $(\sigma_1, \sigma_2) \in [0.5, 2] \times [10^{-4}, 10^{-1}]$. This rectangle is uniformly sampled with $25 \times 25 = 625$ points.

4.3.2. Approximation error convergence

We follow Algorithm 1, initializing the support point set with the four corners of the parameter space, and new support points are added one by one. We also make sure that initial support points are linearly independent.

For each of the 625 sample points, the exact lead field is computed in order to evaluate the ***approximation error*** (see Eq. (4.4)). This is done only for validation purpose and is not necessary for the algorithm. We then compare the approximation error Eq. (4.4) in Fig. 4.2a with its ***upper bound error***, computed with Eq. (4.7). Fig. 4.2a shows that both approximation and upper bound errors ("Approximation 2D" and "Upper bound 2D") decrease exponentially fast with the number of support points. Moreover, they decrease at the same speed. This is an important property because in practice the approximation error is not available, and Fig. 4.2a shows that the decrease of the upper bound error can be used as a convergence criterion for Algorithm 1.

As shown in Fig. 4.2b, with 22 support points, the approximation error from Eq. (4.4) over the sampled conductivity space is bounded by $2 \cdot 10^{-5}$.

We also illustrate our method for a three-dimensional conductivity region, considering the scalp conductivity as an extra parameter. "Upper bound 3D" in Fig. 4.2a shows the decrease of the upper bound error with the number of support points. We can notice that the decrease is still exponential, even if its slope is less steep.

In the case of FEM, the head matrix H is homogeneous with respect to conductivities. It means that up to a constant, manifolds $\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1}$, $\boldsymbol{\sigma} \in \mathbb{R}^2$ and $\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1}$, $\boldsymbol{\sigma} \in \mathbb{R}^3$ are the same. For symmetric BEM, however, $\boldsymbol{H}$ is not homogeneous (even though the lead field is) and so considering scalp conductivity as unknown does increase the dimension of the manifold $\boldsymbol{H}_{\boldsymbol{\sigma}}^{-1}$.

We also evaluated the upper bound error on some conductivity points which did not correspond to any of the samples used in Algorithm 1, and it was not significantly bigger than the ones on sample points. This shows the good continuity properties of our method, as mentioned previously.

4.3.3. Approximation time

The time required for computing 625 exact lead field matrices is ***4 hours***, which amounts to ***23 seconds per matrix*** (2 physical cores @2.60GHz, 16Gb RAM).

Precomputing all required matrices on ***20 support points*** for Algorithm 2 is achieved in ***14 minutes***.

Once these matrices are computed it only takes ***58 seconds*** to approximate the 625 lead fields, which amounts to ***0.09 seconds per matrix***.

As mentioned in the previous section, approximation time does not depend on the complexity of the head model as it only requires to solve a least-squares problem (4.8) on the number of support points. For more complex head models, precomputation time would increase, but the approximation time would remain equal to 0.09 sec for 20 support points (but it may require more support points).

4.3.4. Conductivity estimation with simulated sources

We now examine convergence properties in a realistic application of conductivity estimation.

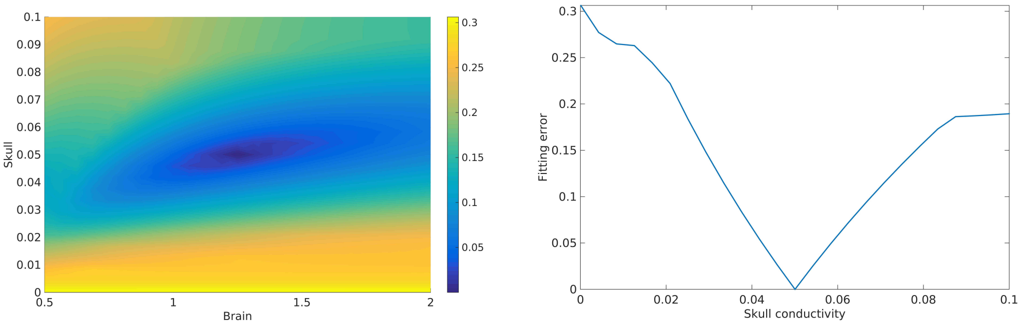
Using the same head model as above, we simulate a dataset $\boldsymbol{y}$ which corresponds to a single dipole source for a reference conductivity (ground truth). Then, we use a simple dipole fitting approach to solve the inverse problem. For each of 625 conductivity samples in the domain, we approximate the lead field matrix and choose the source as the column of this matrix which best fits fits measurement in terms of Euclidean norm. Let $\boldsymbol{M}_j$ represent the j-th column of a matrix $\boldsymbol{M}$. For each conductivity $\boldsymbol{\sigma}$ the data fitting error is defined as follows:

$$R(\boldsymbol{\sigma}) = \min_{j,a} \left\| \boldsymbol{y} - a \cdot \boldsymbol{L}(\boldsymbol{\sigma})_j \right\|_2 \ ,$$

where $\boldsymbol{L}(\boldsymbol{\sigma})_j$ denotes j-th source's lead field, and a - its optimal amplitude. $R(\boldsymbol{\sigma})$ represents the best fitting error of a single dipole corresponding to the measurement $\boldsymbol{y}$ and conductivity $\boldsymbol{\sigma}$.

Computing $R(\boldsymbol{\sigma_i})$ for each of the 625 conductivity samples, we obtain a ***data fitting error map*** on the conductivity domain. The estimated conductivity is the one which minimizes this data fitting error.

Using the exact lead fields, we obtain a data fitting error map (Fig. 4.3a) whose minimum lies at the simulated conductivity point (1.25 for brain and 0.05 for skull). Fig. 4.3b shows a data fitting error for brain conductivity equal to 1.25.



(a) **(b)**

Figure 4.3: Data fitting error map using exact lead fields: (a) full domain of interest. The minimum of error map corresponds to simulated conductivity configuration (1.25 for brain and 0.05 for skull); (b) Fitting error with brain conductivity equal to 1.25 (one column of map a).

We then compute a similar map using approximated lead fields. To show that approximated lead fields are of sufficient quality for conductivity estimation, we compute lead field approximation for a different number of support points. It is visible from Fig. 4.4 that the "shape" of the data fitting error map converges very fast to the exact one. ***10 support points*** are sufficient to correctly estimate the simulated conductivity.

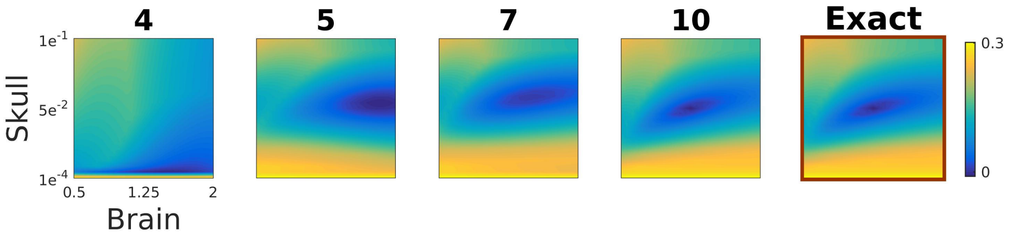
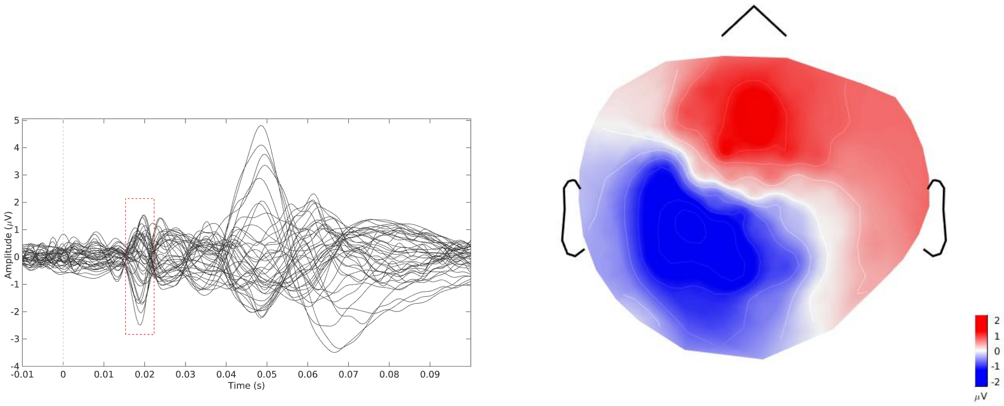


Figure 4.4: Data fitting error map using approximated vs. exact lead fields for simulated data. Number on the top of the maps indicates the number of support points. 10 points are sufficient to obtain an error map similar to the exact one.

4.3.5. Conductivity estimation with real data

We use the same head model and conductivity domain of interest as in the previous section. Real EEG data was taken from a median nerve stimulation experiment. The right median nerve was percutaneously stimulated using monophasic square-wave impulses with a duration of 0.3 ms at 2.8 Hz. As in [Tadel et al., 2016], we filtered and averaged the data, and removed heartbeats and eye movement artifacts using Brainstorm [Tadel et al., 2011].

We are interested in the N20-P20 somatosensory averaged evoked potentials originating from Brodmann's area 3b situated in the posterior bank of the Rolandic fissure [Niedermeyer and Da Silva, 2004, p. 1076]. We can see a remarkable activity peaking at approximately 20 ms which has a dipolar topography on the left hemisphere (Fig. 4.5).



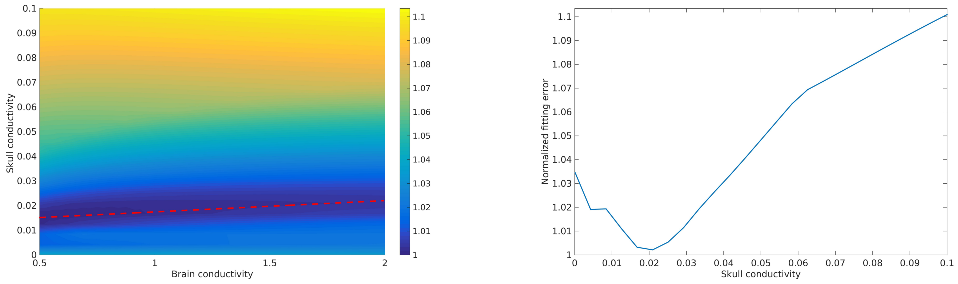
(a) **(b)**

Figure 4.5: Signal of interest: (a) Averaged time series per EEG channel; (b) Topography at 19.5 ms.

We analyzed 5 time samples ($\{\boldsymbol{y}(t)\}_{t=1}^{5}$) of a time window from 0.0185s to 0.0205s, which correspond to the local pick of the signal with dipole-like topography. The data fitting map is computed as follows:

$$R(\boldsymbol{\sigma}) = \min_{j} \sum_{t=1}^{5} \min_{a} \|\boldsymbol{y}(t) - a \cdot \boldsymbol{L}(\boldsymbol{\sigma})_j\|_2 \ .$$

The data fitting error, computed using the exact lead fields, is shown in Fig. 4.6. It is normalized to its minimum value, i.e. $\forall i: \ R(\boldsymbol{\sigma_i}) = \frac{R(\boldsymbol{\sigma_i})}{\min_k R(\boldsymbol{\sigma_k})}$. Many factors contribute to the fitting error: additive noise, wrong conductivity model, simplified head and source models, inverse problem assumption, etc.



(a) **(b)**

Figure 4.6: Data fitting error map using exact lead fields with real data: (a) full domain of interest. The red line represents the optimal skull conductivity for each brain conductivity; (b) Data fitting error with brain conductivity equal to 1.

Because of the different sources of noise, the impact of brain conductivity becomes less important and we cannot significantly find its optimal value. But, for each fixed brain conductivity, we can obtain the skull conductivity which minimizes the error: this skull conductivity value lies between 0.01 S/m and 0.02 S/m. Let us notice that this range is clearly reduced with respect to the range reported in literature. We can also notice that conductivity contributes up to 10% relatively to the other sources of error, which shows the importance of the correct estimation of this parameter.

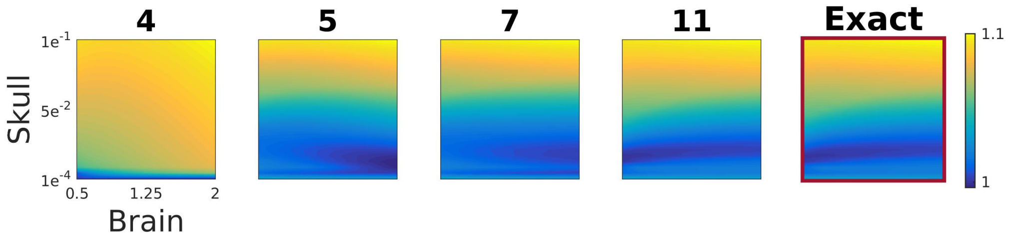


Figure 4.7: Data fitting error map using approximated vs. exact lead fields for real data. Number on the top of the maps indicates the number of support points. 11 points are sufficient to obtain an error map similar to the exact one.

As can be seen in Fig. 4.7, it is enough to use 11 support points to obtain an error map similar to the one obtained using exact lead fields.

4.4. Discussion

Similarly to reduced basis methods used for parametrized PDEs [Hesthaven et al., 2016], our method is based on the assumption that the manifold of parametrized solutions of EEG forward problem can be approximated by low-dimensional linear subspace. As a result, we use similar tools as the reduced basis method, such as support points, upper bound error and greedy basis selection. Our approach, however, is adapted for the particular structure of the EEG forward problem. For instance, we approximate not a single solution of Eq. (2.10), but a family of such solutions, assembled in a so-called lead field matrix: instead of dealing with a reduced basis for Galerkin projections, we deal with the projection of the inverse head matrix onto a precomputed basis.

A different approach of lead field matrix approximation is presented in [Costa et al., 2017], where a polynomial matrix of degree n is used to approximate the exact lead fields. This method has significant differences compared to ours. Firstly, the choice of support points, their number and distribution, is completely arbitrary, while our algorithm includes a basis selection, which is made in an optimal way. Secondly, the polynomial method does not allow to control the approximation error without computing exact solutions. The approach in [Costa et al., 2017] has several advantages, however. To approximate a lead field matrix, it does not need to know its decomposition (i.e. matrices $\boldsymbol{S}$, $\boldsymbol{H_\sigma}$ and $\boldsymbol{D_\sigma}$) nor their dependence on conductivity. Approximated lead fields can be expressed in a closed form as polynomial functions of conductivity, which is easier to work with analytically (e.g. take a gradient) which can also be important for some applications.

4.5. Conclusions

In this work, we introduced a method for fast approximation of the EEG forward solution for different conductivities. Computing the exact lead field for many conductivity configurations is time consuming. Our approach only requires to compute a small number of exact solutions. Under some reasonable assumptions, all other lead fields can be approximated in a fast way with a controlled approximation error. This accelerated computing time allows to explore conductivity space, which is crucial for applications such as head tissue conductivity estimation.

Our method provides both the lead field approximation based on a set of given support points as well as the way to choose this set. The proposed algorithm guarantees the monotonic convergence of the approximation error. Moreover, while the complexity of computing the exact solutions for each support point depends on the number of vertices in a head model, once basis matrices are computed, the complexity of our approximation method is independent of the number of vertices.

Besides the theoretical properties of our method, we studied its empirical performance. Realistic simulations showed an exponential decrease of approximation error with respect to the number of support points. We also demonstrated the usefulness of the method in a realistic context of conductivity estimation from EEG data, both simulated and real. As expected, a relatively small number of precomputed basis matrices provide results which are similar but remarkably faster when compared to using exact matrices.

The main motivation of this work is to propose a tool to boost lead field computation, a necessary step for simultaneous estimation of brain sources' activities and head tissues conductivity. We presented a simple approach for solving this inverse problem, based on single dipole data fitting, to show that our method can be efficiently used for this kind of problem. It would certainly be relevant to consider conductivity estimation with more complex source models (e.g. multiple dipoles), more unknown conductivities (e.g. composite skull structure [Dannhauer et al., 2011]), or different conductivity space sampling approaches. We believe that our method will significantly improve the practical aspects of any such study.

Bibliography

Zeynep Akalin Acar and Scott Makeig. Effects of forward model errors on EEG source localization. *Brain Topography*, 2013. ISSN 08960267. doi: 10.1007/s10548-012-0274-6.

Zeynep Akalin Acar, Can E Acar, and Scott Makeig. Simultaneous head tissue conductivity and EEG source location estimation. *NeuroImage*, 124:168–180, jan 2016. ISSN 10959572. doi: 10.1016/j.neuroimage.2015.08.032.

Facundo Costa, Hadj Batatia, Thomas Oberlin, and Jean-Yves Tourneret. Skull Conductivity Estimation for EEG Source Localization. *IEEE Signal Processing Letters*, 24(4):422–426, apr 2017. ISSN 1070-9908. doi: 10.1109/LSP.2017.2669101. URL `http://ieeexplore.ieee.org/document/7855647/`.

Moritz Dannhauer, Benjamin Lanfer, Carsten H. Wolters, and Thomas R. Knösche. Modeling of the human skull in EEG source analysis. *Human Brain Mapping*, 32(9):1383–1399, sep 2011. ISSN 10659471. doi: 10.1002/hbm.21114. URL `http://doi.wiley.com/10.1002/hbm.21114`.

Sónia I. Gonçalves, Jan C. De Munck, Jeroen P.A. Verbunt, Fetsje Bijma, Rob M. Heethaar, and Fernando Lopes da Silva. In vivo measurement of the brain and skull resistivities using an EIT-based method and realistic models for the head. *IEEE Transactions on Biomedical Engineering*, 50(6):754–767, 2003. ISSN 00189294. doi: 10.1109/TBME.2003.812164.

Alexandre Gramfort, Théodore Papadopoulo, Emmanuel Olivi, and Maureen Clerc. OpenMEEG: opensource software for quasistatic bioelectromagnetics. *BioMedical Engineering OnLine*, 9(1):45, 2010. ISSN 1475-925X. doi: 10.1186/1475-925X-9-45. URL `http://biomedical-engineering-online.biomedcentral.com/articles/10.1186/1475-925X-9-45`.

Jan S Hesthaven, Gianluigi Rozza, and Benjamin Stamm. *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*. SpringerBriefs in Mathematics. Springer International Publishing, Cham, 2016. ISBN 978-3-319-22469-5. doi: 10.1007/978-3-319-22470-1. URL `http://link.springer.com/10.1007/978-3-319-22470-1`.

R. Hoekema, G. H. Wieneke, F. S S Leijten, C. W M Van Veelen, P. C. Van Rijen, G. J M Huiskamp, J. Ansems, and A. C. Van Huffelen. Measurement of the conductivity of skull, temporarily removed during epilepsy surgery. *Brain Topography*, 2003. ISSN 08960267. doi: 10.1023/A:1025606415858.

Jan Kybic, Maureen Clerc, Toufic Abboud, Olivier Faugeras, Renaud Keriven, and Théo Papadopoulo. A common formalism for the Integral formulations of the forward EEG problem. *IEEE Transactions on Medical Imaging*, 24(1):12–28, jan 2005. ISSN 0278-0062. doi: 10.1109/TMI.2004.837363. URL `http://ieeexplore.ieee.org/document/1375158/`.

Seok Lew, Carsten H. Wolters, Alfred Anwander, Scott Makeig, and Rob S. MacLeod. Improved EEG source analysis using low-resolution conductivity estimation in a four-compartment finite element head model. *Human Brain Mapping*, 2009. ISSN 10659471. doi: 10.1002/hbm.20714.

Ernst Niedermeyer and F.H.Lopes Da Silva. *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, Fifth Edition*. Lippincott Williams and Wilkins, 5th edition, 2004. ISBN 0-7817-5126-8.

R Pohlmeier, H Buchner, G Knoll, a Rienäcker, R Beckmann, and J Pesch. The influence of skull-conductivity misspecification on inverse source localization in realistically shaped finite element head models. *Brain topography*, 1997. ISSN 0896-0267. doi: 10.1007/BF01190384.

Stanley Rush and Daniel A. Driscoll. Current Distribution in the Brain From Surface Electrodes. *Anesthesia and Analgesia*, 47:717–723, 1968.

Francois Tadel, Yasuhiro Haruta, Ei-ichi Okumura, and Takashi Asakawa. Yokogawa/KIT tutorial: Median nerve stimulation, 2016. URL `http://neuroimage.usc.edu/brainstorm/Tutorials/Yokogawa`.

Franois Tadel, Sylvain Baillet, John C. Mosher, Dimitrios Pantazis, and Richard M. Leahy. Brainstorm: A user-friendly application for MEG/EEG analysis. *Computational Intelligence and Neuroscience*, 2011:1–13, 2011. ISSN 16875265. doi: 10.1155/2011/879716. URL `http://www.hindawi.com/journals/cin/2011/879716/`.

S Vallaghé and M Clerc. A global sensitivity analysis of three- and four-layer EEG conductivity models. *IEEE Transactions on Biomedical Engineering*, 2008. doi: 10.1109/TBME.2008.2009315.

Sylvain Vallaghé, Maureen Clerc, and Jean Michel Badier. In vivo conductivity estimation using somatosensory evoked potentials and cortical constraint on the source. In *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro - Proceedings*, 2007. ISBN 1424406722. doi: 10.1109/ISBI.2007.357032.

Robert Van Uitert, Chris Johnson, and Leonid Zhukov. Influence of head tissue conductivity in forward and inverse magnetoencephalographic simulations using realistic head models. *IEEE Transactions on Biomedical Engineering*, 2004. ISSN 00189294. doi: 10.1109/TBME.2004.836490.

Gang Wang and Doutian Ren. Effect of brain-to-skull conductivity ratio on EEG source localization accuracy. *BioMed Research International*, 2013. ISSN 23146133. doi: 10.1155/2013/459346.

Chapter 5

Providing a family of solutions to the M/EEG source reconstruction problem through data-driven cortical clustering

Contents

5.1. Overview

In the previous chapter, we saw that most of the M/EEG inverse problem methods are based on convex optimization and, consequently, identify a single solution. Usually, the uniqueness of the solution is considered advantageous and is a desirable property. However, because of the ill-posedness of the problem, it is very likely that other spatially distinct source configurations could explain the data equivalently well as the identified solution. The methods which are forced to find a unique solution, without giving alternative ones, can be disadvantageous, because they may find a false solution. Despite this fact, the problem of finding not one, but several alternative source configurations which can explain the data is not well studied. Some algorithms do estimate not only cortical activity from M/EEG signals, but also analyze the uncertainty of these estimations [Fuchs et al., 2004, Bekhti et al., 2018].

We assume M/EEG forward operator to be computed from a realistic head model. The source space is modeled as a finite set of current dipoles distributed on the cortical surface. We associate one dipolar source to each vertex of the cortical mesh. We assume the orientation of the dipoles to be orthogonal to the cortical surface. The M/EEG forward model is

$$\boldsymbol{Y} = \boldsymbol{L}\boldsymbol{X} + \boldsymbol{B}, \tag{5.1}$$

where $\boldsymbol{Y}$ is the $N_E \times N_T$ matrix representing the signal measured by N_E sensors at N_T time samples, $\boldsymbol{L}$ is the $N_E \times N_S$ lead field matrix whose columns represent lead fields of N_S sources, $\boldsymbol{X}$ is an $N_S \times N_T$ matrix of sources amplitude, and $\boldsymbol{B}$ is an $N_E \times N_T$ additive noise matrix. For the rest of the chapter, we consider white Gaussian noise. If, in practice, it is not the case, whitening should be applied [Engemann and Gramfort, 2015]. The inverse problem aims at finding $\boldsymbol{X}$ knowing $\boldsymbol{Y}$ and $\boldsymbol{L}$.

In this chapter, we consider brain activity to be localized within a set of cortical regions. We assume each region to be connected, i.e. that each pair of vertices can be connected by a path within the region. We also assume that each dipole in the active region has the same amplitude.

We propose a new approach for M/EEG source reconstruction, which is able to find several candidate solutions and that we call CLUstering Based MUSIC or CLUB-MUSIC. These candidates are different in term of their sizes and/or positions on the cortex, but fit the data with similar accuracy. In sections 5.2 – 5.5, we introduce a method based on the agglomerative hierarchical clustering [Murtagh and Contreras, 2012] whose objective is not to reconstruct a single active cortical region but to find several candidates, which similarly fit the M/EEG data. In Section 5.6, we propose a method to estimate the number of active sources.

We summarize our approach in Section 5.7. We will provide numerical results on simulated and real data in the next chapter.

5.2. General description of the clustering algorithm

Agglomerative hierarchical clustering consists in producing hierarchical representations in which the clusters at each level of the hierarchy are created by merging clusters at the lower level, based on the pairwise dissimilarities between clusters [Hastie et al., 2009].

We define a cluster c_i as a set of vertices of a cortical mesh and denote its size by $|c_i|$. $\boldsymbol{L}(c_i)$ denotes the lead field vector (also called ***topography***) of a cluster, which, because of the constant source hypothesis, is a sum of the lead field vectors of its dipole sources (corresponding columns of the matrix $\boldsymbol{L}$). For any pair of clusters c_i and c_j, we define a dissimilarity measure that we will call ***clustering error***:

$$E(i,j) = F(i,j,\boldsymbol{Y}) + R(i,j) \tag{5.2}$$

The first term $F(i,j,\boldsymbol{Y})$ represents the data-related error that would be obtained if the clusters are merged (see Section 5.4). $R(i,j)$ represents a regularization term which will be discussed in Section 5.5.

We represent the neighborhood information between clusters as a function $N(i,j) = N(j,i) = 1$, if clusters c_i and c_j are neighbors and 0 otherwise. Two vertices are neighbors if they share an edge on the cortical mesh. Given M/EEG data $\boldsymbol{Y}$, a cortical mesh and a lead field matrix $\boldsymbol{L}$, the clustering algorithm proceeds with the following steps:

Step 1. Initialize clusters considering each vertex of the cortical mesh (and associated dipole source) as a cluster. Initialize neighborhood $N(i,j)$ based on the cortical mesh. Initialize A as the set of current cluster indices.

Step 2. Examine all ***inter-neighbor*** clustering errors (see Eq. (5.2)) and merge the two clusters which minimize it:

$$i^*, j^* = \underset{i,j \in A;\ N(i,j)=1}{\operatorname{argmin}} E(i,j);\ c_k = c_{i^*} \cup c_{j^*}$$

where k is the index of the new cluster.

Step 3. Compute the lead field of the new cluster: $\boldsymbol{L}(c_k) = \boldsymbol{L}(c_{i^*}) + \boldsymbol{L}(c_{j^*})$

Step 4. Replace the two merged clusters by the new cluster and update neighborhood information: $A = A \setminus \{i^*, j^*\} \cup \{k\}$

$$\forall i \in A : N(i,k) = \begin{cases} 1, & \text{if } N(i,i^*) = 1 \text{ or } N(i,j^*) = 1 \\ 0, & \text{otherwise} \end{cases}$$

Step 5. Return to Step 2 and repeat until the whole cortex is one cluster.

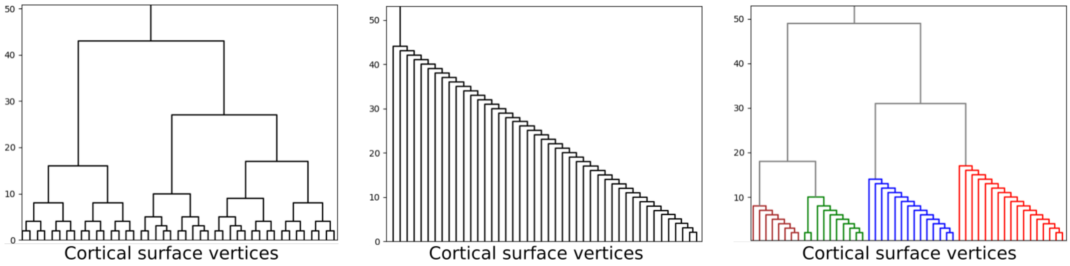
Merging two clusters can be seen as growing one region in the direction that locally minimizes the regularized data fitting error. Taking into account the neighborhood information guarantees connected regions. The way the lead field for new clusters is computed constrains these regions to have constant activity, i.e. for all dipoles of an active region to have the same amplitude.

To represent the result of an hierarchical agglomerative clustering, a binary tree, so-called ***dendrogram*** can be plotted. It provides a highly interpretable complete description of the clustering in a graphical format [Hastie et al., 2009].

5.3. Dendrogram analysis

In this section, we introduce the general framework for working with dendrograms, obtained after performing clustering algorithm described in the previous section. We do not yet propose the exact formulas for the error and the regularization terms F and R from Eq. (5.2). They will be discussed in the following sections.

We represent the dendrogram, using the size of the clusters as a y-axis, which is monotonically increasing. Fig. 5.1 shows some different types of possible dendrograms.



(a) (b) (c)

Figure 5.1: Different types of the dendrogram that can be obtained after clustering of the cortex. (a) This dendrogram is characterised by systematically merging clusters of about the same size. (b) This dendrogram is characterised by increasing the size of a single cluster by merging it with a small number of vertices at a time. (c) This dendrogram is a combination of previous two types – it consists of a set of "smoothly growing" clusters which, at some point, start to merge with each other.

The dendrogram, represented in Fig. 5.1a, is characterised by systematically merging clusters of about the same size. In contrast, the dendrogram in Fig. 5.1b can be interpreted as a single "smoothly growing" cluster. The dendrogram, shown in Fig. 5.1c is a combination of both these types – it consists of a set of "smoothly growing" clusters which, at some point, start to merge with each other. This type of dendrogram is of particular interest, because it provides us with a useful and intuitive feature – ***a growing region*** object.

Before giving the definition of this object, let us first remind some important terminology involved when working with dendrograms. First of all, a ***node*** in the dendrogram corresponds to each cluster, obtained after applying a clustering algorithm. If cluster c_k is obtained by merging clusters c_i and c_j, the i-th and j-th nodes in the dendrogram, those are called ***children*** or ***descendants*** of the ***parent*** k-th node. The initial clusters (cortical mesh vertices, in our case) are called ***leaves***.

A ***sub-tree*** of a dendrogram is a tree consisting of a node in this dendrogram and all of its descendants.

A growing region *is a sub-tree of a dendrogram, such that:*

1. *For each of its nodes, the size of at least one of its 2 children is smaller than some value* ***s****.*

2. *This sub-tree is not nested into another sub-tree, which also verifies the first property.*

To extract growing regions from a dendrogram, we should find nodes where sub-trees stop growing "smoothly". We will call these nodes ***cutting points***. Based on the definition of a growing region, starting from the bottom of the dendrogram, we check nodes: if $c_k = c_i \cup c_j$ and $|c_i| \geq |c_j|$, then c_i is a cutting point if $|c_j| > s$, where s is an arbitrarily chosen merging "speed" threshold. For example, with $s = 2$, the dendrogram in Fig. 5.1c can be decomposed into four growing regions (marked in different colors). See Fig. 5.2 for more details.

In the context of the M/EEG-based cortical clustering for source reconstruction, we are interested to obtain a dendrogram, which can be decomposed into the a of growing regions. We will see in following sections that, with the proper choice of the clustering error (Eq. (5.2)), this will be the case.

Here are some important properties of such dendrograms and its growing regions:

- A data-related error F is associated to each node of the dendrogram.

- A lead field vector (topography) is associated to each node of the dendrogram.

- Each node of the dendrogram represents a connected cortical region.

- Nodes of a growing region represent nested cortical regions (Fig. 5.2).

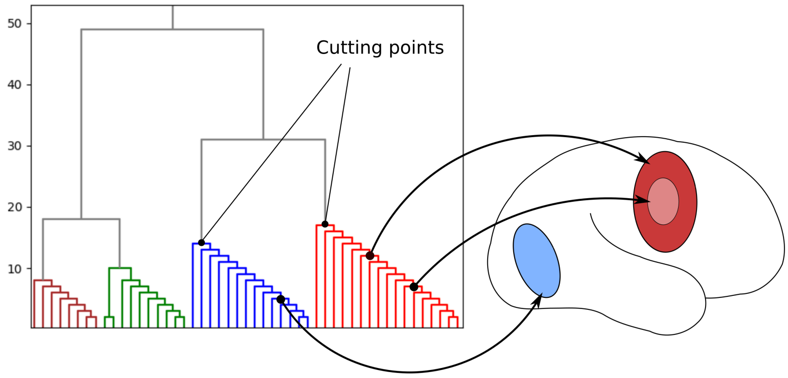


Figure 5.2: Colored sub-trees of the dendrogram correspond to four growing regions. Nodes of a growing region correspond to nested cortical regions (shades of red on the cortex). Nodes of different growing regions correspond to spatially separated cortical regions (blue vs. red on the cortex).

- Nodes from the different growing regions represent spatially non-overlapping cortical regions (Fig. 5.2).

There is an error value associated to each cortical region, which shows how the topography generated by this region explains the measurements. It allows us to find the node within each growing region which has the minimum associated error. After finding such a "best node" for each of growing regions, we may then determine which growing regions best explain the data.

Let us remark that for each leaf of a growing region, there is a unique path which relates it with the root (cutting point) of the growing region. It is due to the fact that each node has a single parent. Consequently, for a given node, the number of paths going through it equals to the number of leafs, corresponding to the node. Based on these observations, we give following definition:

A growing path *of a growing region is the longest path relating a leaf and the root of the growing region and containing its best node. The length of a path is the number of the nodes it contains.*

Extracting a growing path from a growing region allows us to analyse how the error F evolves as a function of a cortical regions size (Fig. 5.3). Taking the longest possible path passing through the best node, allows us to have the smallest possible step size.

Analysing the growing path of a growing region allows us not only to determine the regions' optimal size, but also to estimate possible size bounds, or how sensitive the error is with respect to the regions size.

To conclude, we are interested in obtaining dendrograms of the type presented in Fig. 5.1c, which can be decomposed into a set of growing regions. Based on the

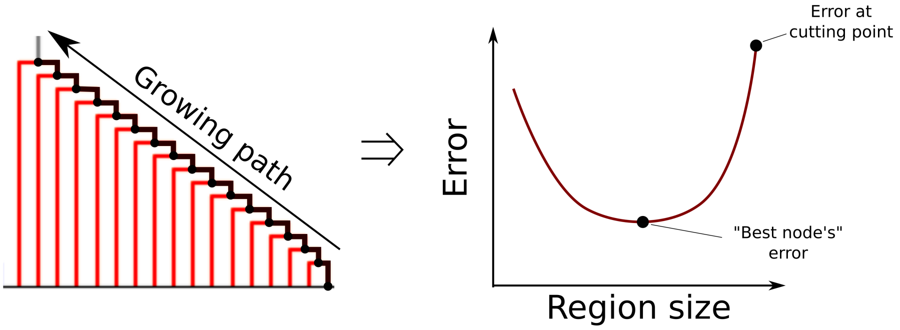


Figure 5.3: Example of a growing region (red dendrogram) and its growing path (black line). There are a size and an error associated to each node of the growing path. It allows to represent the error as a function of the regions size.

"best nodes'" error, we can select the best candidate growing regions for source reconstruction : growing regions with the lowest best error. By extracting the growing path for each candidate, we may also estimate the size bounds for each candidate growing region.

Naturally, the type of dendrogram that we obtain after clustering strongly depends on the choice of the clustering error (Eq. (5.2)). In Sections 5.4 and 5.5, we will propose the accuracy and regularization terms F and R which are relevant to constraints discussed in Section 5.1, and which provide dendrograms decomposable into growing regions.

5.4. Data fitting term

We propose to use a data fitting term based on the MUSIC criterion. As described in Chapter 3, the MUSIC approach is based on computing the projection of a candidate topography $\boldsymbol{l}$ onto the signal subsapce $\boldsymbol{P}_s$, and then selecting the topography which maximizes this value:

$$\mu(\boldsymbol{l}) = \frac{||\boldsymbol{P}_s\boldsymbol{l}||^2}{||\boldsymbol{l}||^2}.$$

Let us notice that maximizing MUSIC localizer $\mu(\boldsymbol{l})$ is equivalent to minimizing the following value:

$$\operatorname*{argmax}_{\boldsymbol{l}} \mu(\boldsymbol{l}) = \operatorname*{argmin}_{\boldsymbol{l}} \frac{||\boldsymbol{l} - \boldsymbol{P}_s\boldsymbol{l}||^2}{||\boldsymbol{l}||^2} = \operatorname*{argmin}_{\boldsymbol{l}} \frac{||(\boldsymbol{I} - \boldsymbol{P}_s)\boldsymbol{l}||^2}{||\boldsymbol{l}||^2}.$$

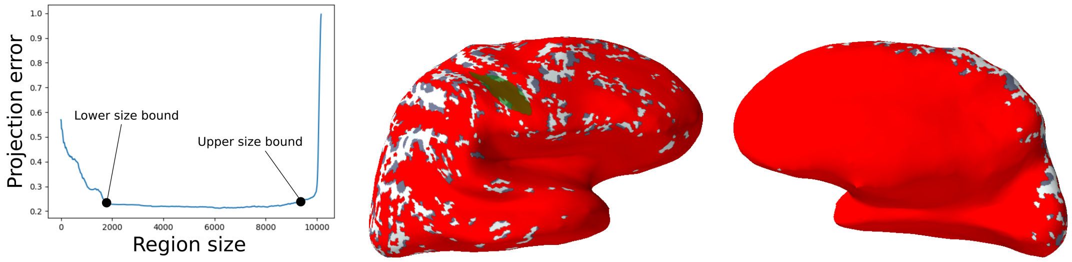
Based on this property, we propose the following, MUSIC-based data fitting component for our clustering algorithm:

$$F(i,j,\boldsymbol{Y}) = \frac{\|(\boldsymbol{I}-\boldsymbol{P}_s)(\boldsymbol{L}(c_i)+\boldsymbol{L}(c_j))\|^2}{\|\boldsymbol{L}(c_i)+\boldsymbol{L}(c_j)\|^2}, \tag{5.3}$$

where $\boldsymbol{L}(c_i)+\boldsymbol{L}(c_j)$ is the topography obtained by merging clusters c_i and c_j, and $\boldsymbol{P}_s$ is a signal subspace of measurements $\boldsymbol{Y}$. In this and the following chapter, we will call this value $F(i,j,\boldsymbol{Y})$ a MUSIC-like projection error or just a ***projection error***.

5.5. Regularization term

Without regularization ($R(i,j) = 0$), we face an overfitting problem. After simulating a single active cortical region (Fig. 5.4) and applying our clustering algorithm, the obtained dendrogram of the right hemisphere is of the type presented in Fig. 5.1b – a single region is growing by merging a small amount of vertices at each step. As a result, the whole dendrogram is a single growing region. If we extract its growing path, we can analyze how the projection error is evolving with the size of the region (Fig. 5.4a). We might notice that the error decreases at the beginning and then stays at the same level (between the lower and upper size bounds points on the plot) until at the end where it rapidly increases. This is mostly due to merging dipoles whose lead fields cancel each other out, and by merging "silent" sources - dipoles with a small lead field norms.



(a) (b) (c)

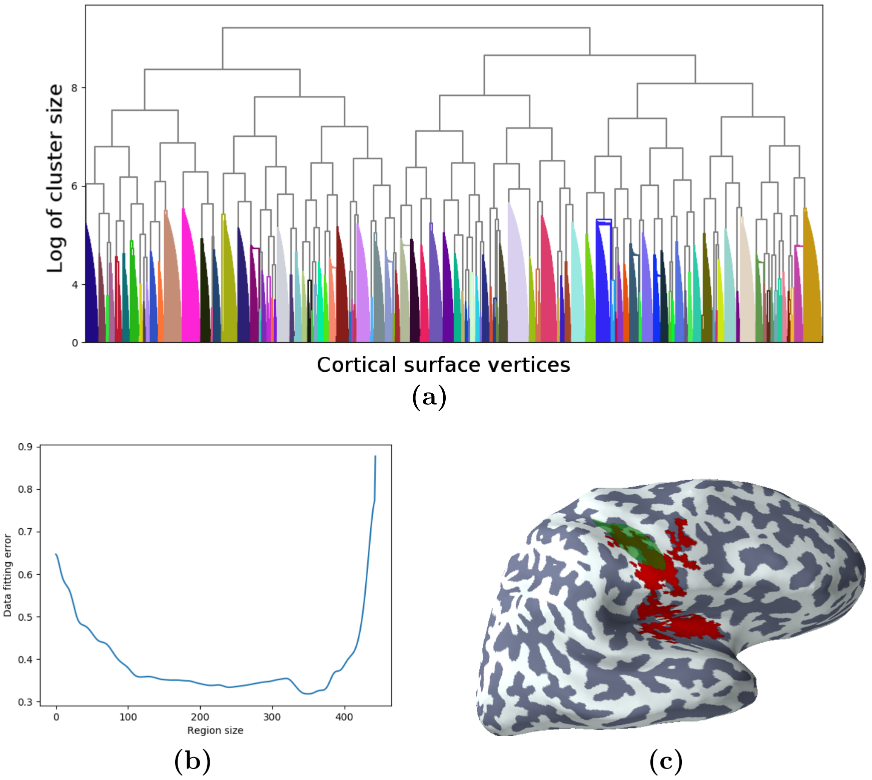
Figure 5.4: Example of applying a clustering algorithm without regularization. One cortical region was simulated (in green). After applying clustering algorithm we obtain a single growing region which covers the whole hemisphere. (a) The projection error of the region as a function of size. After some point (lower size bound) the error remains constant, while the region keeps growing. (b) The cortical region (in red) corresponding to the upper size bound of the growing region, lateral view. (c) The upper size bound region, medial view.

To solve this problem, a first approach that comes to mind is to penalize the size of regions (to prevent them from having too big size). This can be done by introducing a regularization term from Eq. (5.4) in the clustering error from Eq. (5.2):

$$R(i,j) = \lambda \cdot (|c_i| + |c_j|)^2, \tag{5.4}$$

where $\lambda > 0$ is a regularization hyperparameter which controls the importance of the penalty term, and $|c_i|$ denotes the size of the cluster c_i. The choice of the exponent 2 in the size regularization term is arbitrary and will be discussed later.

After applying the clustering algorithm with the size regularization term from Eq. (5.4), we obtain a dendrogram that can be decomposed into a set of growing regions (Fig. 5.5a).



(a)

(b)

(c)

Figure 5.5: Example of applying a clustering algorithm with size regularization. One cortical region was simulated (in green). (a) Obtained dendrogram, which can be decomposed into growing regions (in color). (b) The data-related error of a region as a function of size. (c) The region (in red) with the optimal size (minimizing data-related error).

Based on the best node's error of each growing region, we may select the best candidates for source reconstruction – growing regions with the minimum projection error. In our case, the second best candidate happened to be the nearest to the simulated region in terms of the cortical location. Fig. 5.5b shows the projection error as a function of the region size. First, the error decreases, then reaches the plateau where it varies very little and at the end it starts to increase.

Fig. 5.5c shows the cortical region, corresponding to the best node of the candidate growing region. We may notice that it has an odd, unrealistic shape.

To constrain the shape of the reconstructed region, another regularization term can be proposed, which penalizes regions with irregular borders.

First, for two clusters c_i and c_j, we define a value $B(i,j)$ as the number of vertices in c_i which have at least one neighbor in c_j. Thus, the value $B(i,j)+B(j,i)$ can be interpreted as a mutual length of the border between clusters c_i and c_j.

We can define the following objectives for the regularization:

1. Regions should have smooth borders. It can be achieved by maximizing the inter-cluster border length $B(i,j)+B(j,i)$ (or minimizing $\frac{1}{B(i,j)+B(j,i)}$). Let us consider an example – based on the maximization of the border length. point 1 in Fig. 5.6a is more favorable than point 2 to be merged with the cluster of black points, because it has a longer "merging border".

2. Regions should grow smoothly, i.e. by merging it a small number of vertices at a time. It can be achieved by minimizing the size of the smallest of two merging clusters (minimizing the term $\min(|c_i|, |c_j|)$). It will help to obtain a dendrogram, which could be decomposed into a set of growing regions.

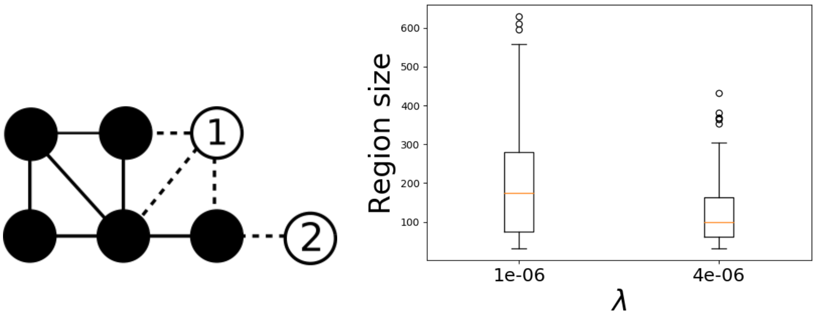
3. The shape regularization should be less important than projection error for small regions, becoming more important while a region is growing. It can be achieved by adding term $\lambda(|c_i| + |c_j|)^2$ to the minimization problem. This term is small for small regions, letting them grow freely at the beginning, and becomes more important as the regions grow larger. The hyperparameter λ defines how fast regularization becomes important.

We thus propose the following regularization term for minimization, which is the product of these three terms:

$$R(i,j) = \lambda \cdot (|c_i| + |c_j|)^2 \cdot \min(|c_i|, |c_j|) \cdot \frac{1}{B(i,j) + B(j,i)} \tag{5.5}$$

Note: The power 2 for the size component $(|c_i| + |c_j|)^2$ has been chosen empirically. The regularization would also be efficient with a exponent equal to 1 or 3, but we found that 2 gives the best compromise between the projection error and shape regularization.

The regularization parameter λ controls the distribution of the growing region sizes - higher value results in smaller regions (Fig.5.6b). Thus, for the rest of this chapter, we will use the clustering error based on the projection error (Eq. (5.3)) and the regularization term (Eq. (5.5)):



(a) **(b)**

Figure 5.6: (a) Points 1 and 2 are two candidates for merging with the cluster of black points c_i. For Point 1 $B(i,1) + B(1,i) = 3 + 1 = 4$ and for Point 2 it is only 2. Thus merging point 1 gives a smaller regularization value than merging point 2. (b) The distribution of the maximum size of growing regions (the size of the cutting point nodes) for different regularization parameter.

$$E(i,j) = \frac{\|\big(\boldsymbol{I} - \boldsymbol{P}_s\big)\big(\boldsymbol{L}(c_i) + \boldsymbol{L}(c_j)\big)\|^2}{\|\boldsymbol{L}(c_i) + \boldsymbol{L}(c_j)\|^2} + \lambda(|c_i| + |c_j|)^2 \frac{\min(|c_i|, |c_j|)}{B(i,j) + B(j,i)}. \tag{5.6}$$

5.6. Source selection

We so far provided an algorithm to decompose the cortex into a set of growing regions, each characterized by its growing path, a MUSIC-like projection error and topography associated to each node. Now we define the source selection problem whose objectives are:

1. Determine the number of real distinct cortical sources.
2. Associate each growing region to one of these sources, i.e. provide candidates for each source).

In Chapter 3, we discussed source selection with the RAP-MUSIC method. It is able to select source topographies but it does not use the idea of several candidates for a source. Instead of reworking the RAP-MUSIC algorithm, we propose another, more general approach to source selection, based on finding local minima of the projection error. The local minima are defined not in the cortex space, but in the topographies space.

We start by selecting the topographies corresponding to the best nodes of each growing region extracted from the dendrogram. We then normalize these topographies by dividing them by their norms and assemble them into a matrix $\hat{\boldsymbol{L}}$. We can interpret columns of $\hat{\boldsymbol{L}}$ as samples of an N_E-dimensional unit sphere. Based on this representation, we can define the distance between two unit vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ as $d(\boldsymbol{x},\boldsymbol{y}) = \arccos(|\langle \boldsymbol{x},\boldsymbol{y}\rangle|)$. Notice the use of the absolute value of the dot product, because, the sign of the lead field vector does not matter. This value represents the angle between two vectors, and thus may be converted to degrees.

Assume also that a function f is defined on columns of $\hat{\boldsymbol{L}}$, i.e. that a value $f(\hat{\boldsymbol{L}}_i) \in \mathbb{R}$ is associated to the i-th column of $\hat{\boldsymbol{L}}$. Based on the angular distance on the space of normalized lead fields, we define a local minimum as:

Definition 3. *$\hat{\boldsymbol{L}}_i$ is a α-**local minimum** of f if:*

$$\forall j \textit{ such that } d(\hat{\boldsymbol{L}}_i, \hat{\boldsymbol{L}}_i) \leq \alpha,\ f(\hat{\boldsymbol{L}}_i) \leq f(\hat{\boldsymbol{L}}_i).$$

Taking $f(\hat{\boldsymbol{L}}_i) = \|(\boldsymbol{I} - \boldsymbol{P}_s)\hat{\boldsymbol{L}}_i)\|$, where $\boldsymbol{P}_s$ is the projector onto the signal subspace (which, in fact, is the projection error used for clustering), we formulate the problem of selection of the source regions as finding α-local minima of f for a fixed α, which we call ***a neighborhood angle***.

The angular distance is a natural way of measuring how similar two topographies are. We want to extract different topographical components that explain the signal and then for each of them region sources that can explain those topographies. Let us notice that two spatially distinct cortical regions may have similar topographies and thus will be considered as neighbors in terms of angular distance. And, conversely, two regions can be near each other on the cortex but have distinct topographies. In this case, these regions are not considered as neighbors.A topography which is selected as a local minimum will be called a ***source topography***.

We may represent the information about selected source topographies for different neighborhood angle values as shown in Fig. 5.7. For each angle, we plot the projection error value, associated to the selected topographies (black points on the graph). It gives us information about the number of selected topographies as well as about their accuracy in terms of the projection onto the signal subspace. Naturally, the number of local minima increases when the angle becomes smaller. This representation can be a powerful and intuitive tool for source topographies selection. In fact, well chosen angle and error thresholds define the number of possible topographies. We can see in Fig. 5.7 that for angles smaller than 20 degrees (blue line represents an angle threshold), we start to get a lot of local minima. We may also notice that for angles between 20 and 40 degrees, first three sources have significantly lower error than the remaining two sources. This gives an idea about the possible error threshold (red line represents an error threshold). So, from this figure, we can infer that, most likely, there are three active sources.

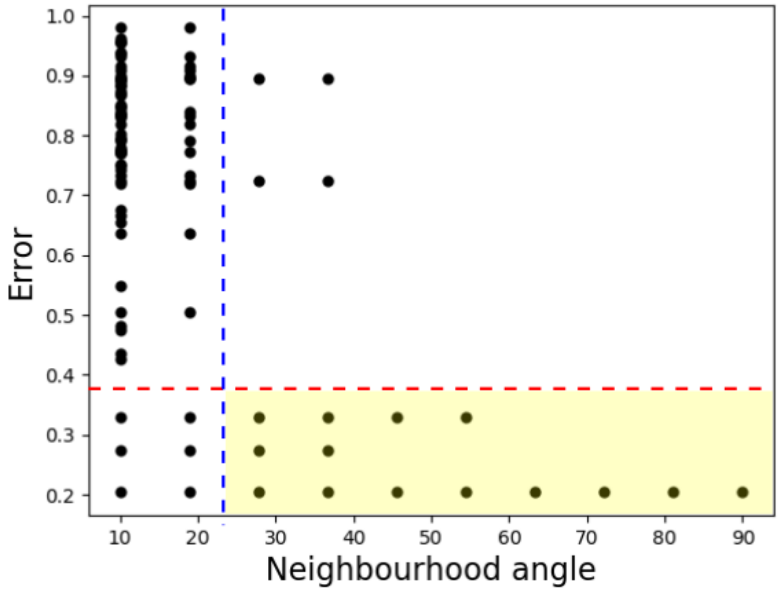


Figure 5.7: Projection errors corresponding to the local minima with respect to the neighborhood angle. Blue line is the angle threshold, red line is the error threshold. Three sources seem to be a reasonable choice for selection.

Once the local minima topographies are selected for a fixed angle, we may associate each growing region to one of the selected source topographies. Thus, for each selected topography, we may form a list of associated ***candidate*** growing regions, based on the nearest angular neighbor search. This allows to separately analyse each source topography in terms of its best candidate source region.

5.7. Summary of the CLUB-MUSIC algorithm

Here are the principal steps of the CLUB-MUSIC algorithm. Examples of its implementation can be found in the next chapter.

1. Perform SVD of the data and select its first singular values as a basis of the signal subspace.
2. Apply algorithm presented in Section 5.2 to construct a dendrogram.
3. Cut the obtained dendrogram into a set of growing regions, as described in Section 5.3.
4. Perform the source topographies selection as described in Section 5.6. For each selected topography, we obtain a set of candidate growing regions.
5. The best candidates for each source topography can be analysed in terms of their projection error, size and location on the cortex (see examples in the next chapter).

5.8. Conclusions

In this chapter, we introduced a new method of M/EEG source reconstruction. Here are some important points, concerning our method:

- It is based on the assumption that the M/EEG data is generated by the unsynchronized activity of cortical regions.
- The key feature of our method is so-called ***growing region object***.
- Our method provides not a single, but several spatially distinct candidates for a solution, which explain the data with similar accuracy.
- With our approach, it is possible to estimate not only the position of the candidate region, but also its spatial extension.

In the nest chapter we will test our method with simulated and real M/EEG data.

Bibliography

Yousra Bekhti, Felix Lucka, Joseph Salmon, and Alexandre Gramfort. A hierarchical bayesian perspective on majorization-minimization for non-convex sparse regression: application to m/eeg source imaging. *Inverse Problems*, 2018.

Denis A Engemann and Alexandre Gramfort. Automated model selection in covariance estimation and spatial whitening of MEG and EEG signals. *NeuroImage*, 108:328–342, 2015. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2014.12.040. URL `http://www.sciencedirect.com/science/article/pii/S1053811914010325`.

Manfred Fuchs, Michael Wagner, and Jörn Kastner. Confidence limits of dipole source reconstruction results. *Clinical Neurophysiology*, 2004. ISSN 13882457. doi: 10.1016/j.clinph.2004.01.019.

T. Hastie, R. Tibshirani, and J.H. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer series in statistics. Springer, 2009. ISBN 9780387848846. URL `https://books.google.fr/books?id=eBSgoAEACAAJ`.

Fionn Murtagh and Pedro Contreras. Algorithms for hierarchical clustering: An overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2012. ISSN 19424787. doi: 10.1002/widm.53.

Chapter 6

Results of CLUB-MUSIC on simulated and real MEG data

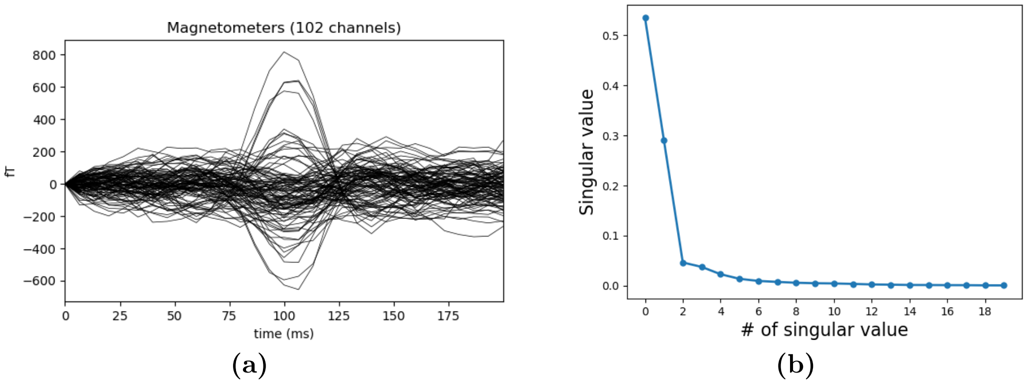
Contents

6.1. Overview

In this chapter we test our approach on simulated and real data. In Section 6.2 we start with simulating a single active region to demonstrate in detail the type of analysis that can be provided with the principal feature of our method – growing region object. Then, in Sections 6.3 and 6.4, we simulate multiple active cortical regions, to test how our method deals with source selection for different levels of SNR. In Section 6.5, we perform source reconstruction with our method on a real auditory evoked MEG data set.

6.2. Simulated MEG data: single active region

We use the "Sample" subject from MNE-python software data set [Gramfort et al., 2013] to simulate the data. The source space is represented as the vertices of a cortical mesh with an order of 10000 vertices per hemisphere (totalling 20470 vertices). Dipole orientations are fixed to be orthogonal to the cortical surface. The lead field matrix is computed with a three-layered head model (brain, skull, scalp) using the standard BEM method, implemented in MNE-python. We simulate a connected cortical source region with uniform amplitude. We simulate a MEG signal (102 magnetometer channels) and add a multivariate spatially uncorrelated and temporally IIR-filtered Gaussian noise, provided by MNE-python tools (Fig. 6.1a).



(a) **(b)**

Figure 6.1: (a) Simulated MEG signal with additive noise (b) Squares of the singular values of the normalized signal.

Based on the singular value decomposition (SVD) of the simulated signal (Fig. 6.1b), we choose its first three singular vectors to form a 3-dimensional projection signal subspace. We then apply the clustering algorithm, presented in the previous chapter, to obtain a dendrogram.

Note: The value of the regularization parameter λ is set empirically to $4 \cdot 10^{-6}$ and we use the same value in all the experiments on simulations or real data. The proper way of setting its value is an open question. As for the parameter s for cutting the dendrogram, we used $s = 10$, but the precise value of this parameter is not important, as long as it is relatively small. We use the same value for parameter s in all the experiments. Clustering and growing regions extraction took ***5 min***. The computation time, however, can be potentially reduced. First, two hemispheres are processed independently, and thus the computation can be parallelized. Moreover, the python code we used can also be optimized to significantly reduce the computation time.

After cutting the dendrogram as described in the previous chapter, we obtain a set of sub-trees, that we call growing region objects. Fig. 6.2 shows the obtained dendrogram with growing regions marked in different colors. Each node of such a sub-tree represents a connected cortical region of certain size with associated lead field vector (topography) and projection error.

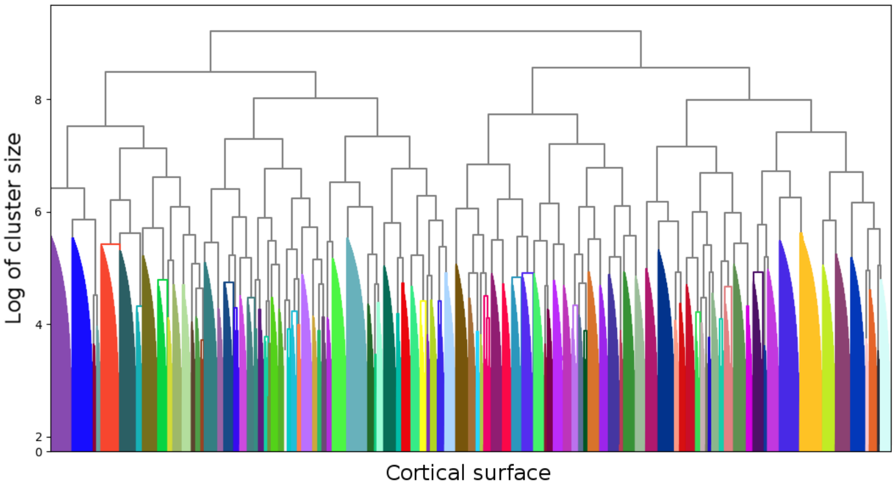
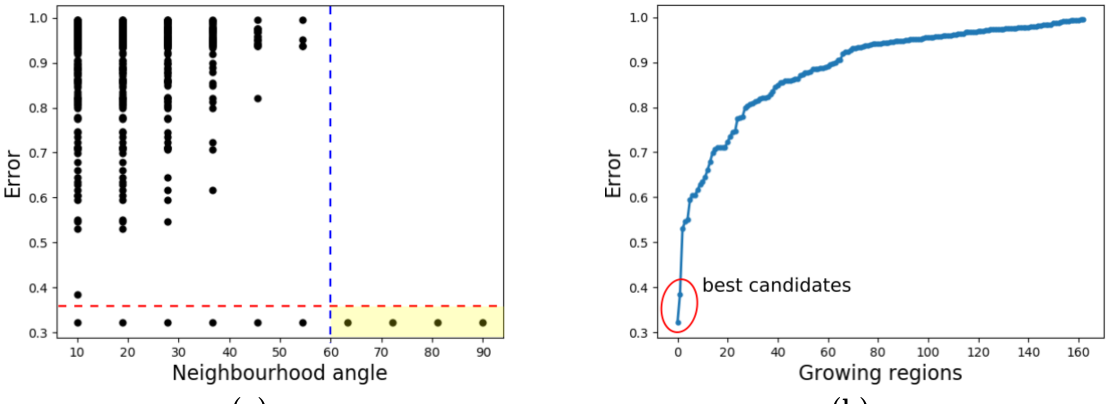


Figure 6.2: Dendrogram, obtained by reconstruction a single simulated region with CLUB-MUSIC. It can be decomposed into growing regions (in colors).

For each growing region, we find the node which has the lowest projection error. We assemble associated topographies into a matrix $\hat{\boldsymbol{L}}$ to perform a source selection, as described in Section 5.6 of the previous chapter. We use a range of 10 neighborhood angles from 10 to 90 degrees. For each such a neighborhood angle, we extract the topographies, corresponding to the local minima of the projection error, so-called ***source topographies***. Fig. 6.3a shows the errors associated to local minima with respect to the neighborhood angle. We may notice that the smaller the angle, the more are local minima. Well chosen angle and error thresholds define the number of possible source topographies. We can see in Fig. 6.3a that for angles smaller than 60 degrees (the blue line representing an angle threshold), there are many local minima with significantly higher error than the global minimum (the red line representing an error threshold). From this figure, it is clear that there is most likely a single active source, which indeed corresponds to the simulation. Let us notice that, for the case of multiple active sources, this representation remains useful, even though the selection of thresholds becomes less obvious (see next sections).

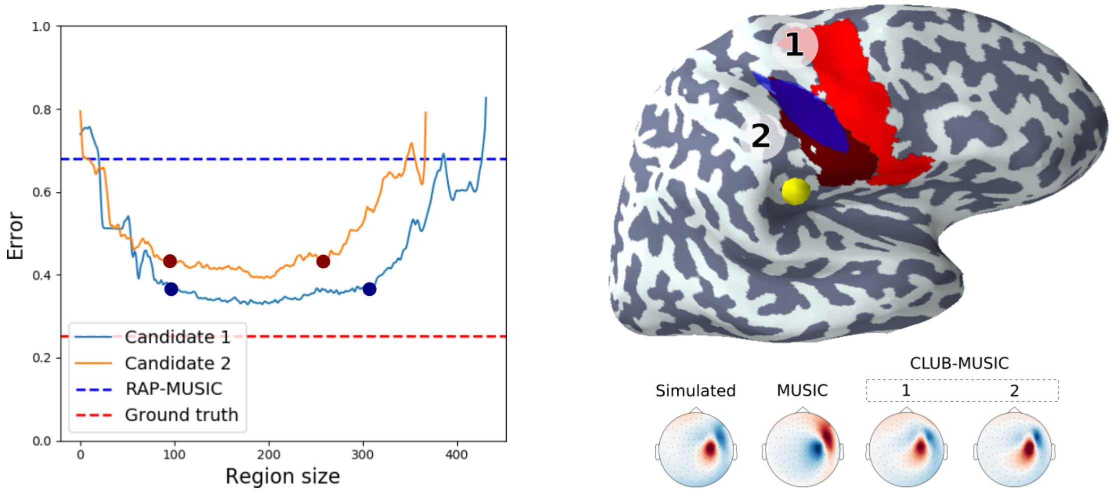


(a) **(b)**

Figure 6.3: (a) Projection errors corresponding to the local minima with respect to the neighborhood angle. Blue line is the angle threshold, red line is the error threshold. A single source topography is selected, which corresponds to the simulation (b) Best node projection errors corresponding to the candidate growing regions. Two candidates have significantly lower error than other candidates.

Fig. 6.3b shows the best projection error associated to each growing region. We may notice that two regions have significantly lower error than others. We select them as main candidate regions for the source reconstruction. For each candidate growing region, we can now observe how the projection error depends on the size of the region. For this, we select a growing path for each candidate and plot how the error evolves along this path (Fig. 6.4a). For both selected candidates, we see that the error significantly decreases at the beginning, then it stays relatively stable and finally starts to increase when regions become larger and the regularization term (Eq. (3.3)) starts dominating.

By choseing an appropriate threshold (Fig. 6.5a), we can estimate the lower and upper bounds of region extensions. The lower bound is the point, when the curve crosses the threshold for the first time. The upper bound is the point, when the curve crosses the threshold for the last time. The choice of the threshold is arbitrary. We use a value equal to *minimum error* $+ 0.05$ for each candidate in all the experiments in this chapter.



(a) (b)

Figure 6.4: (a) Projection error as a function of the region size for two selected candidates. Blue line corresponds to the projection error obtained with the dipole, found by MUSIC. Red line represents the true level of noise. Dots on the curves correspond to the lower and upper size bounds for two candidates. (b) Top: cortical representation of the simulated region (in blue) and the upper bound of two candidates (shades of red). Yellow marker represents the location of the dipole found by MUSIC. Bottom: topographies of the simulated source and those found by MUSIC and our method.

(a) (b) (c)

Figure 6.5: (a) Arbitrary chosen threshold (red line) determines regions lower and upper size bounds (red dots). The threshold value we used is *minimum error* + 0.05. (b) Cortical representation of the lower (in yellow) and upper (in red) region bounds. (c) Thresholded sLORETA solution at the peak of the signal.

Fig. 6.4b shows cortical regions, representing the upper bounds of two candidates (marked in shades of red). The blue region is the simulated one. As we can see, our algorithm is able to closely reconstruct the simulated region (candidate

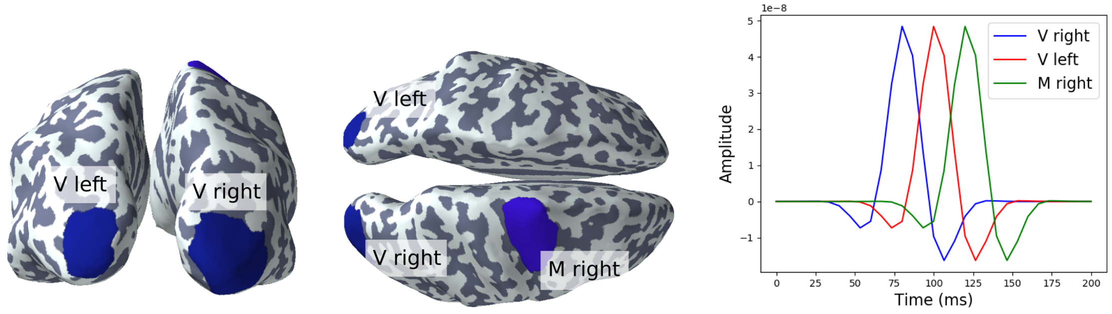
2), as well as provide another candidate (candidate 1). Let us notice that the candidate with the lowest projection error (candidate 1) is not the region that we simulated. It confirms our remark that the M/EEG is so ill-posed that several spatially separated cortical regions can explain the data with similar accuracy. This is also evident from topographies (Fig. 6.4b) produced by the best nodes of each candidate – they are similar to each other as well as to the simulated one.

Fig. 6.4b also shows the dipole location found by standard MUSIC algorithm, performed on the same data with the same signal subspace. We used the free dipole orientation for this method, to get the best data fitting error. The reconstructed dipole is located quite far from the simulated region and has a projection error significantly higher than the one obtained with our method (Fig. 6.4a). Using constrained dipole orientation does not improve the localization.

Fig. 6.5b shows the lower (in yellow) and the upper bound (in red) regions of candidate 2, with the error threshold presented in Fig. 6.5a. Fig. 6.5c shows the reconstruction with the sLORETA method. As we can notice, sLORETA is not able to provide a good spatial information in the case of the focal cortical activity that we simulated.

6.3. Simulated data: multiple active regions with high SNR

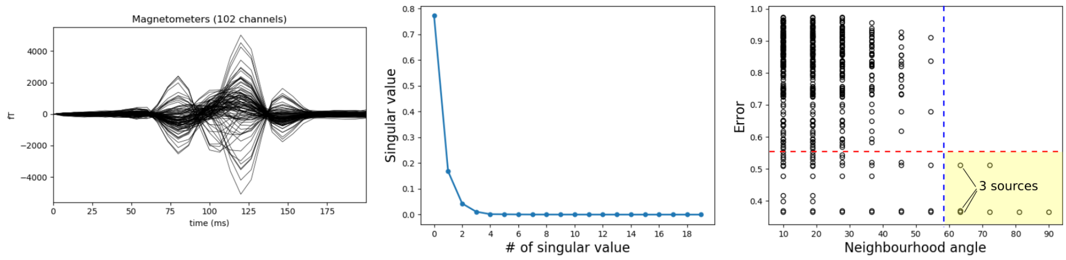
After testing our approach with a single simulated region, we now consider multiple active regions. As for the single region case, we use the "Sample" subject from the MNE-python. We simulate the activity of three separate cortical regions: primary visual cortex of the right and the left hemispheres and primary motor cortex of the right hemisphere. We first simulate cortical regions inside the left and right lateral-occipital and the right precentral parcels of the Desikan-Killiany atlas [Desikan et al., 2006] (Fig. 6.6a). Each simulated region produces the same activity, but with a 20ms delay (Fig. 6.6b).



(a) (b)

Figure 6.6: (a) Simulated cortical regions. V left and V right are regions of visual cortex and M right is a region of motor cortex. (b) Simulated waveforms of each region.

We then apply a forward operator to simulate evoked MEG data. A low-amplitude spatially uncorrelated IIR-filtered Gaussian noise is added to the measurements. Fig. 6.7a shows the simulated signal. The data SNR is 16.5 dB.

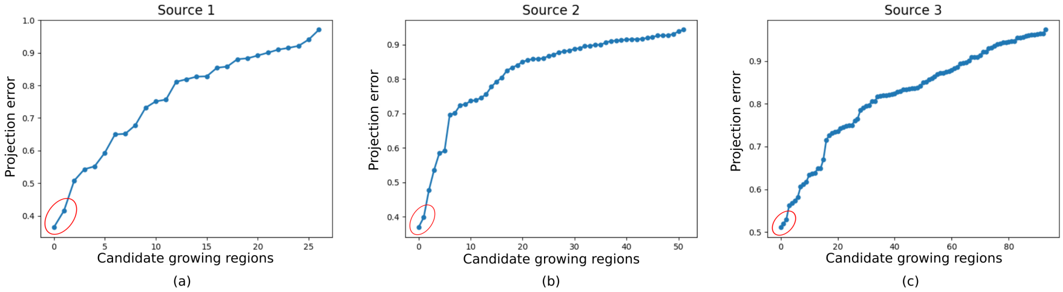


(a) (b) (c)

Figure 6.7: (a) Simulated signal at the sensors level. (b) SVD of the MEG signal. 4 first singular vectors are chosen as a basis of the signal subspace. (c) Projecion errors associated to local minima with respect to the neighborhood angles. 3 source topographies seem to be a reasonable choice for the selection. Two of the minima have almost the same error, so they are hard to visually distinguish as the two points on the graph overlap.

6.3.1. Source localization with CLUB-MUSIC

Based on the the singular value decomposition of the data (Fig. 6.7b), we choose 4-dimensional signal subspace for projection. After applying the CLUB-MUSIC algorithm and decomposing the dendrogram into the growing regions, we may see in Fig. 6.7c the projection errors associated to the local minima with respect to the neighborhood angle. With reasonable angle and error thresholds, we can conclude that there are 3 potential source topographies, which corresponds well to the simulation. We will now separately analyse each of the selected source topographies.



(a) (b) (c)

Figure 6.8: The data fitting error associated to the best nodes of candidate growing regions for each of three selected source topographies. (a) We select first to growing regions as candidates for the Source 1, because they have significantly lower projection error. (b) First two candidate growing regions are selected for the Source 2. (c) First three candidate growing regions are selected for the Source 3.

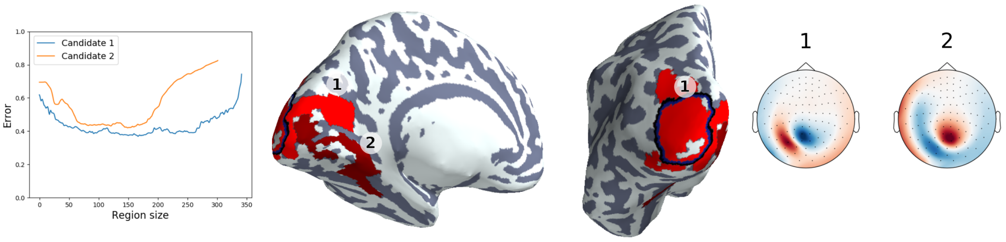
Source topography 1. Looking at the projection error associated to each candidate growing region for source 1 (Fig. 6.8a), we can notice that the first 2 candidates have remarkably lower error compared to the other candidates. Fig. 6.9a shows how the error depends on the size for these 2 candidates. The upper bound candidate regions are plotted on the cortex in Fig. 6.9b. We may see that the first candidate corresponds well to the simulated region on the left visual cortex. The second candidate is a region, near to the simulated one, but located mostly on the Lingual gyrus. The difference in projection error is however very small, and the topographies of two candidates are quite similar up to the sign (Fig. 6.9c).

Note: what matters is the "form" of a topography, not its sign. For example, two candidate topographies in Fig. 6.9c have different sign (because of the opposite region surface orientations) but have similar "form".

Source topography 2. Like for the first source, looking at the projection error associated to each candidate growing region for source 2 (Fig. 6.8b), we can notice that the first 2 candidates have remarkably lower error. Fig. 6.10a shows how the error depends on the size for these 2 candidates. The upper bound candidate regions are plotted on the cortex in Fig. 6.10b. In this case, the candidate 1 which is numerically the best one, does not correspond the to the simulated region on the right visual cortex. The candidate 2 however corresponds well to the simulated region. Fig. 6.10c shows how similar the candidate topographies are.

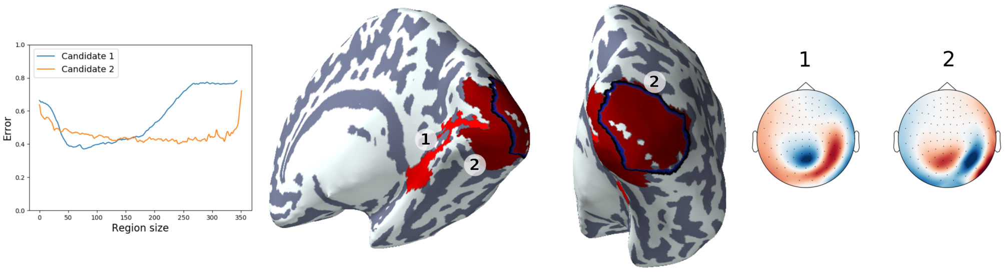
Source topography 3. Three best candidate growing regions for source 3 have similar projection error (Fig. 6.8c). Fig. 6.11a shows how the error depends on the size for these 3 candidates. We may notice that the error of candidate 3 is unstable and the region stays small until it merges with other regions. Candidates 1 and 2, however, seem stable and have similar projection error during the whole growing path. But what is remarkable, is their location on the cortex (Fig. 6.11b). Candidate 1 correspond well to the simulated region on the right motor cortex.

Candidate 2, however, is located far from it. Look at Fig. 6.11c, we may see that first two candidates have quite different topographies. The second candidate appears to be impacted by the topographies generated by the second and third sources. Its location on the cortex follows the same principle – it is located in between the two corresponding simulated regions.



(a) (b) (c)

Figure 6.9: Source 1. (a) Candidate projection error as a function of a region size. (b) Candidate upper bound regions on the cortex (shades of red) and simulated region (dark border). (c) Topographies associated to the best nodes of candidate growing regions.



(a) (b) (c)

Figure 6.10: Source 2. (a) Candidate projection error as a function of a region size. (b) Candidate upper bound regions on the cortex (shades of red) and simulated region (dark border). (c) Topographies associated to the best nodes of candidate growing regions.

6.3.2. Comparison to RAP-MUSIC

We also reconstruct the sources with RAP-MUSIC for comparison with our approach. We use exactly the same data and, to be consistent, we use the same dimension of the signal subspace (4 first singular vectors). The dipole orientation is set free, to get the best possible data fitting error. After applying RAP-MUSIC method with MNE-Python, we select the first 3 found sources for analysis.

Fig. 6.12 shows the locations of the dipoles reconstructed by RAP-MUSIC on the cortex, compared to the regions, found by our method. The RAP-MUSIC dipoles are located on the border or even outside the simulated regions.

(a) (b) (c)

Figure 6.11: Source 3. (a) Candidate projection error as a function of a region size. (b) Candidate upper bound regions on the cortex (shades of red) and simulated region (dark border). (c) Topographies associated to the best nodes of candidate growing regions.

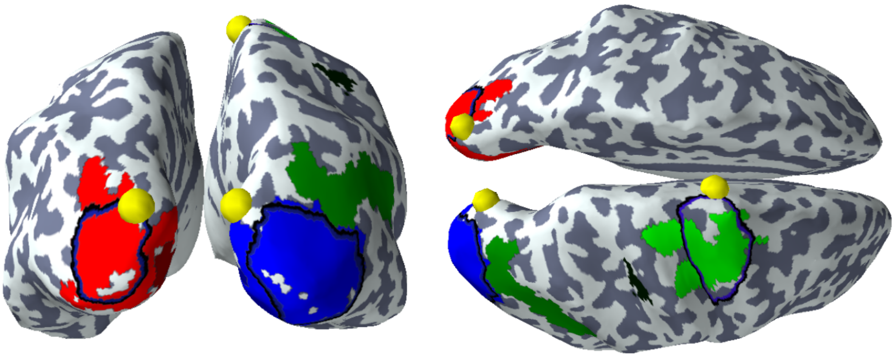


Figure 6.12: Cortical representation of the RAP-MUSIC solution (yellow markers) on top of the regions, found by our method (red, green and blue regions). Different colors correspond to different sources. Shades correspond to different candidates. Lighter shade means lower projection error.

To compare the results of the RAP-MUSIC with our method in terms of the data fitting error, we need to restrict all $2+2+3=7$ candidates found by CLUB-MUSIC into a triplet of sources. We take all possible triplets of candidates for each source (total of $2\times2\times3=12$ triplets). For each such i-th triplet, we form a lead field matrix $\boldsymbol{L}_i$ with three columns, and compute source and the sensor signal reconstruction as follows:

$$\widehat{\boldsymbol{S}}_i = \boldsymbol{L}_i^{\dagger}\boldsymbol{Y}; \quad \widehat{\boldsymbol{Y}}_i = \boldsymbol{L}_i\widehat{\boldsymbol{S}}_i.$$

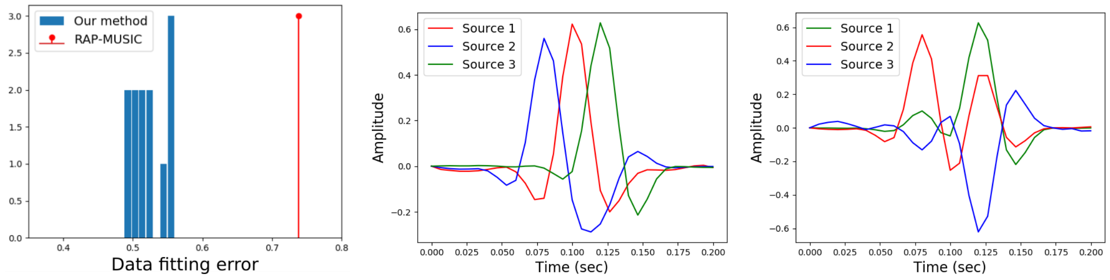
where $\widehat{\boldsymbol{S}}$ is a source estimate, $\widehat{\boldsymbol{Y}}$ is a reconstructed sensor signal and symbol † represents pseudo inverse.

For each triplet, we compute a data fitting error $E_i = \dfrac{\|\boldsymbol{Y}-\widehat{\boldsymbol{Y}}_i\|}{\|\boldsymbol{Y}\|}$ and select the triplet which minimizes it. Fig. 6.13a shows the histogram of the data fitting errors corresponding to the triplets compared to the error obtained with RAP-MUSIC.

We may notice that the error does not vary a lot for different triplets, which confirms our thesis that all candidates for each source can explain the data with similar accuracy. Moreover, each triplet explain data better than he RAP-MUSIC solution.

Fig. 6.13b and Fig. 6.13c show the source signals reconstructed with the best triplet found by our method and RAP-MUSIC respectively. We may see that our method provides waveforms similar to the simulated ones, while RAP-MUSIC does not – peacks of waveforms do not have the same latency as simulated.

Fig. 6.14 shows the sensor signals reconstructed with our method and RAP-MUSIC. In addition to quantitative results, visual analysis also shows signal that we obtained with the best triplet from our method is closer to the original one.



(a) **(b)** **(c)**

Figure 6.13: (a) Distribution of triplets' data fitting error (blue histogram) and the error obtained with RAP-MUSIC (red indicator). (b) Source signals found with our method. (c) Source signal found with RAP-MUSIC.

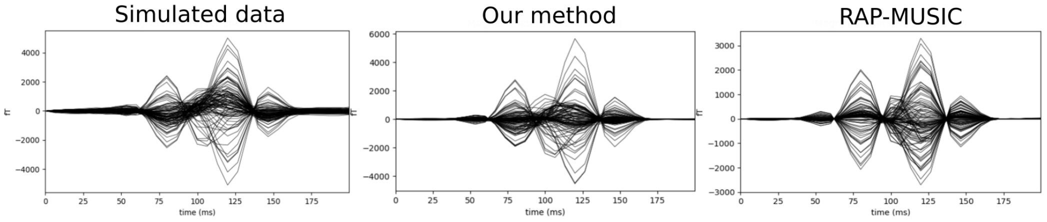
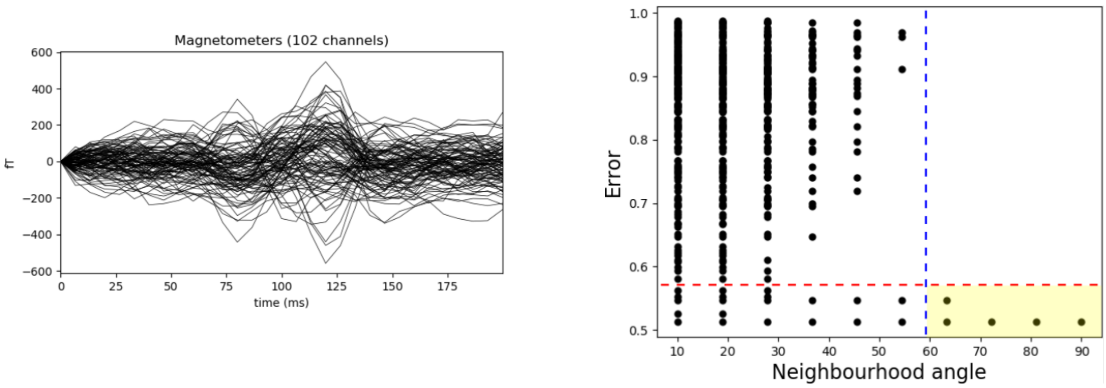


Figure 6.14: The simulated signal at sensor level and the signal reconstructed with our and RAP-MUSIC methods.



(a) **(b)**

Figure 6.15: (a) Simulated MEG data, multiple cortical regions, low SNR. (b) Local minima projection error by neighborhood angle. Two source topographies are selected for further analysis.

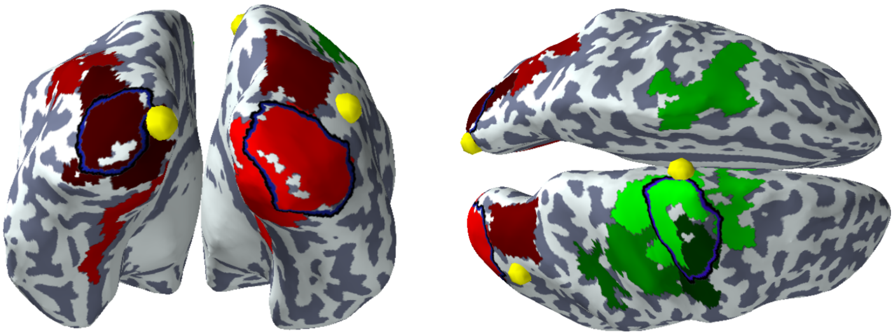


Figure 6.16: Cortical representation of the simulated regions (dark borders) and reconstructions with our method (regions in color) and with RAP-MUSIC (yellow markers) for low SNR data. Color shades correspond to different candidates. Lighter shade means lower projection error.

6.4. Simulated data: multiple active regions with low SNR

In this section, we apply our method to the same data but with much a lower SNR, which now equal to 0.15 dB. Let us notice that we used exactly the same noise as in the case of high SNR, but with a lower amplitude. Fig. 6.15a shows the simulated signal at the sensor level. A 6-dimensional signal subspace is chosen for projection, based on the SVD of the signal.

After applying our method and cutting the dendrogram into a set of growing regions, we represent local minima errors with respect to the neighborhood angles (Fig. 6.15b). Only two sources are selected in this case. In fact, because of the low SNR, left and right visual areas were considered as a single source (in terms of their topographies) (Fig. 6.16).

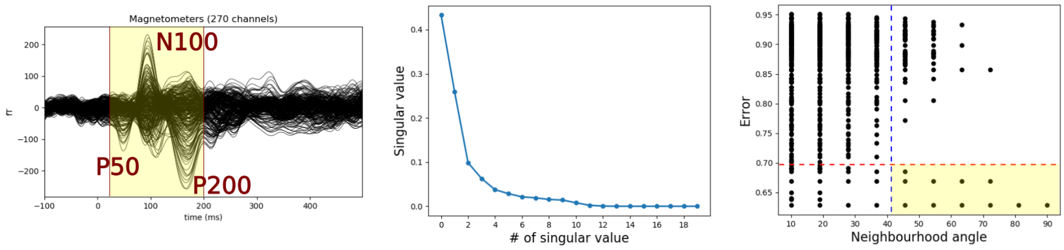
RAP-MUSIC is able to reconstruct 3 sources, but as with the case of high SNR, reconstructed dipoles are located on the border or outside the simulated regions.

6.5. Results on real MEG data

For testing our algorithm with real data, we use the MEG auditory dataset from Brainstorm software which is documented and freely available for download under the GNU general public license [Tadel et al., 2011]. We would like to acknowledge the authors Elizabeth Bock, Peter Donhauser, Francois Tadel and Sylvain Baillet for this dataset. We process the data with the MNE-Python software. You may find the exact pre-processing pipeline on the corresponding MNE tutorial page. The only difference is that in this work we use $0.01 - 30$ Hz frequency band for filtering, instead of $0-100$ Hz as indicated in the tutorial. Here is a brief summery of the experiment, data and pre-processing. Subject was stimulated binaurally with intra-aural earphones (air tubes+transducers). The data, corresponding to 200 regular beeps (440Hz) is considered here. MEG is acquired at 2400Hz, with a CTF 275 system with the subject in sitting position. The data was recorded at the Montreal Neurological Institute in December 2013. The data from 270 MEG axial gradiometers were considered here. The 1.5T MRI of the subject was processed with FreeSurfer 5.3. A set of bad segments was identified and used to reject epochs that overlap with them. The saccades were removed by using SSP [Uusitalo and Ilmoniemi, 1997]. The raw signal was filtered with $0.01 - 30$ Hz frequency band. Epochs were defined as a [-100ms, 500ms] window around a stimulus. Only the first 40 epochs from each run are averaged to obtain evoked data. Data and lead field whitening was performed. Noise covariance is estimated with empty room recordings, available in the dataset.

Fig. 6.17a shows the evoked signal obtained after pre-processing and averaging 40 epochs containing auditory event-related field (ERF). We can identify three main components of the wave:

- *The P50 component.* It has been argued that the P50 reflects pre-attentive arousal due to the appearance of a new event in the auditory scene [Remijn et al., 2014].
- *The N100 component.* This component is generated by multiple sources including the superior temporal plane, the superior temporal gyrus, and perhaps even frontal motor cortex [Niedermeyer and Da Silva, 2004, Woods, 1995].
- *The N200 component.* Located on the supratemporal cortex, this component



(a) (b) (c)

Figure 6.17: (a) Average signal representing the response to a simple auditory stimulation, with main wave components (P50, N100, P200). Window of the signal of interest from 30ms to 200ms is marked with yellow. (b) SVD of the signal of interest. 10 first singular vectors are chosen as a basis of the signal space. (c) Errors associated to the selected source topographies with respect to the used neighborhood angle. 3 topographies appear to be a reasonable choice for the selection.

is generated by a source slightly distinct from the N100 [Papanicolaou et al., 1990].

For our study, we consider the interval between 30ms and 200ms as a signal of interest (yellow area in Fig. 6.17a).

6.5.1. Source localization with CLUB-MUSIC

Based on the singular value decomposition of the signal of interest (Fig. 6.17b), a subspace of dimension 10 is chosen to represent the signal space. The CLUB-MUSIC algorithm is applied to extract growing regions. Clustering and growing regions extraction took ***5 min***. Similarly to the the case of simulated data, we analyze the number of local minima in topography space and associated errors as a function of a neighborhood angle (Fig. 6.17c). The choice of 3 source topographies appears to be reasonable. Let us analyze each of selected source topographies.

Source topography 1. Looking at the projection error associated to each of the candidate growing regions for source 1 (Fig. 6.18a), we can see that the first 4 candidates have remarkably lower error. Fig. 6.18b shows how the error depends on the region sizes for first 4 candidate growing regions. The upper bounds of the candidate regions are plotted on the cortex in Fig. 6.18c. Selected candidates are located on the left hemisphere on the auditory cortex as well as on adjacent supramarginal gyrus and even postcentral gyrus. This means that the data can be explained with each of these regions separately as well as with their simultaneous activity. Fig. 6.18d shows the topographies of these candidates, which are quite similar.

Source topography 2. As with source 1, looking at the error associated to each of the candidate growing regions for the source 2 (Fig. 6.19a), we can select 4 candidates with remarkably lower errors. Fig. 6.19b shows how the error depends on the region sizes. The upper bounds of the candidate regions are plotted on the cortex on Fig. 6.19c. Selected candidates are located on the right hemisphere mostly on the frontal cortex. The best (in terms of data fitting) candidate is located on the Brodmann area 44 (pars opercularis). This area is known to be related to the language and music processing Brown et al. [2006]. Other regions are mostly located on the inferior part of the precentral gyrus. Fig. 6.19d shows the topographies of these candidates which are also quite similar. The data can be explained with each of these regions separately as well as with their simultaneous activity without significant difference in data fitting error.

Source topography 3. For this source, we identify 5 candidate growing regions to analyse. In contrast to the previous two sources, the selection of this number of candidates is less obvious because the error increases more smoothly with respect to candidates (Fig. 6.20a). Fig. 6.20b shows how the error depends on the region sizes. We can see in Fig. 6.20c that selected candidates are mostly located on the superior temporal gyrus and supramarginal gyrus, which corresponds to the auditory cortex. Fig. 6.20d shows the corresponding topographies. Even though there is certain variebility in it, all of them explain the total data with similar accuracy and it is impossible to conclude (within the framework of our method) which of them are the true sources.

6.5.2. Comparison to RAP-MUSIC

We also estimate the sources with RAP-MUSIC method for comparison with our approach. To be consistent, we use exactly the same data, the dimension of the signal space (10 first singular vectors) and the noise covariance. The dipole orientation is free. We use the save noise covariance as for our method. After applying RAP-MUSIC method with MNE-Python, we select the first 3 reconstructed sources.

Fig. 6.21 shows the locations of found RAP-MUSIC sources on the cortex, compared to the candidate regions found by our method. We see that our method provides more spatial information about possible source localization. To compare the results of RAP-MUSIC with our method in terms of the data fitting error, we need to restrict all $4 + 4 + 5 = 13$ candidates into 3 sources. We take all possible triplets of candidates for each source (total of $4 \times 4 \times 5 = 80$ triplets), for each such i-th triplet we form a lead field matrix $\boldsymbol{L}_i$ with three columns, and compute the source and the sensor signal reconstruction as we did for simulated data. For each triplet we compute the data fitting error $E_i = \dfrac{\|\boldsymbol{Y} - \widehat{\boldsymbol{Y}}_i\|}{\|\boldsymbol{Y}\|}$ and select the triplet which minimizes it. Fig. 6.22a shows the histogram of the data fitting error corresponding to the different triplets compared to the error obtained with RAP-MUSIC. We may notice that the error varies very little for different triplets, which confirms our hypothesis that all candidates for each source can explain the data with similar accuracy. Moreover, each triplet explain the data better than the RAP-MUSIC solution.

Fig. 6.22b and Fig. 6.22c show the source signals reconstructed with our method using the best triplet and RAP-MUSIC respectively. What is remarkable here is that time curses, reconstructed with our method are less correlated then the signal reconstructed with RAP-MUSIC.

Fig. 6.23 shows the sensor signals reconstructed with our method and RAP-MUSIC. Visual analysis also shows that the signal that we obtained with our method is closer to the original one.

To conclude, CLUB-MUSIC reconstructed cortical regions, which are known to be involved in auditory processing. Waveforms, reconstructed with CLUB-MUSIC are less correlated than those of RAP-MUSIC and it is possible to relate their peaks to P50, N100 and P200 components, presented in MEG measurements.

6.6. Conclusions

In this chapter, we tested our source reconstruction method on simulated and real MEG data. It confirmed our observation that because of the ill-posedness of the M/EEG inverse problem, several distinct cortical regions may explain the same data with similar accuracy. Thus, even in the simple case of a single simulated region, our method found out several good candidates for reconstruction. In the case of multiple sources and low SNR, the uncertainty becomes even higher. However, our method has also shown a good performance in these situations, as it is grounded on the principle that it is better to obtain false positive regions, than a single false negative. It also outperformed the state-of-art RAP-MUSIC method in terms of the data fitting error.

Several aspects of our method need further analysis and discussion. The most important of them are the choice of the regularization parameter and thresholds for source selection.

Being based on the MUSIC method, our method also shares some of its limitations. For instance it is not designed to work with spatially separated sources which are perfectly synchronized.

We think that despite its limitations, the innovative features that our method provides will make it useful to the M/EEG community.

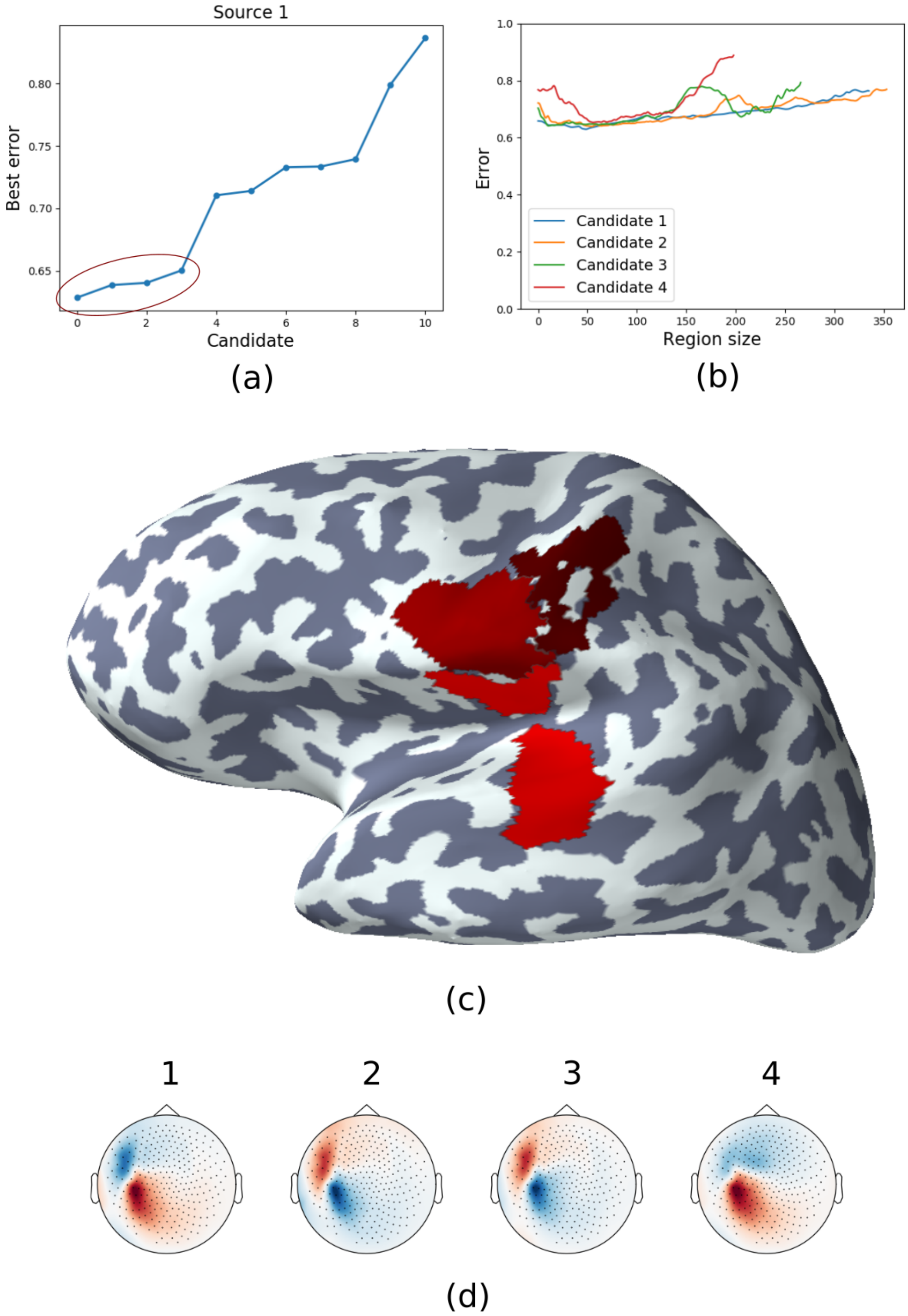


Figure 6.18: Source 1. (a) Best projection error by candidate. First 4 candidates have significantly lower error and are chosen for further analysis. (b) Error as a function of a region size for selected candidates. (c) Upper bound of candidates location on the cortex. Lighter is the shade of red, lower is the error of corresponding region. (d) Topographies associated to the best nodes of candidate growing regions.

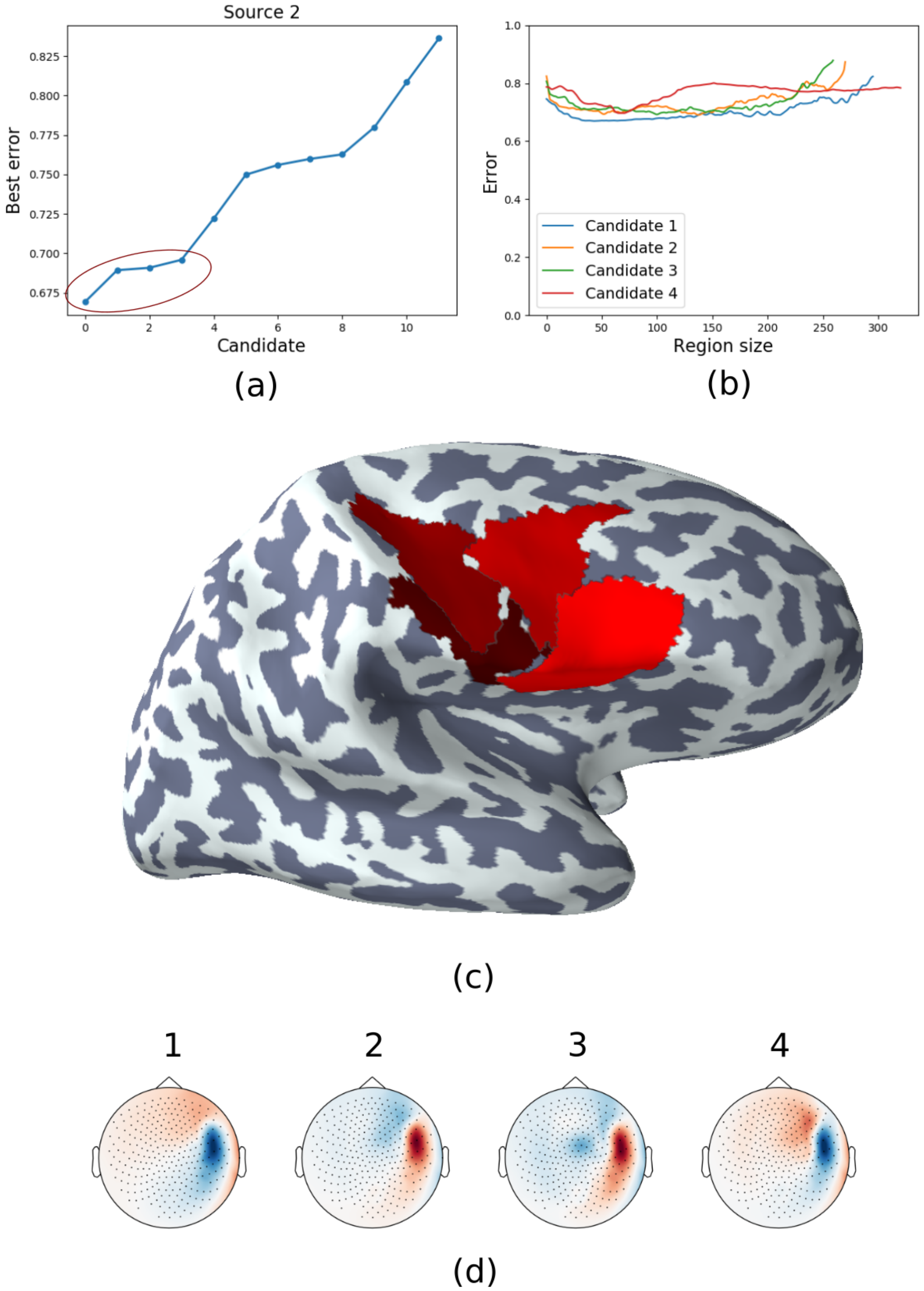
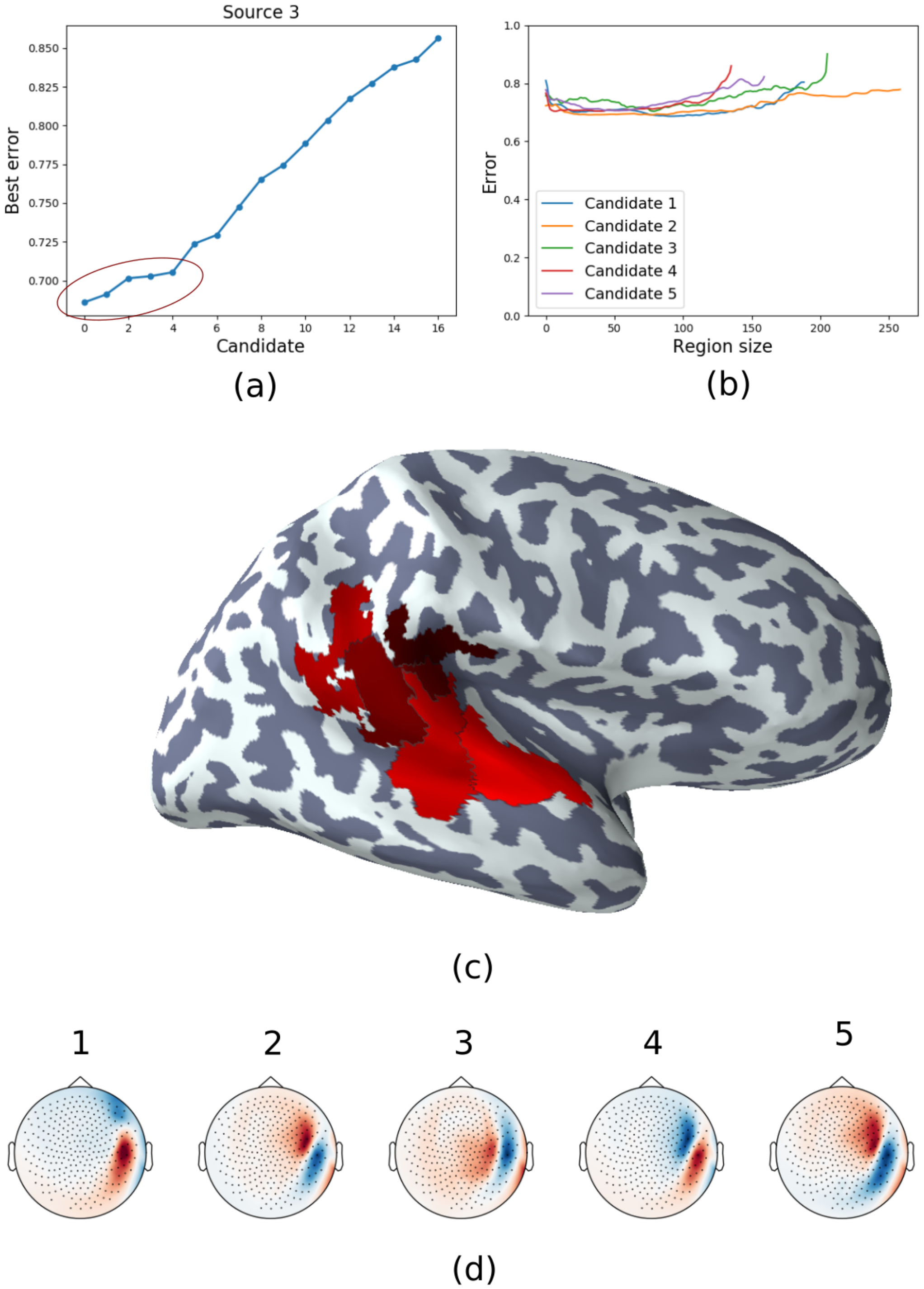


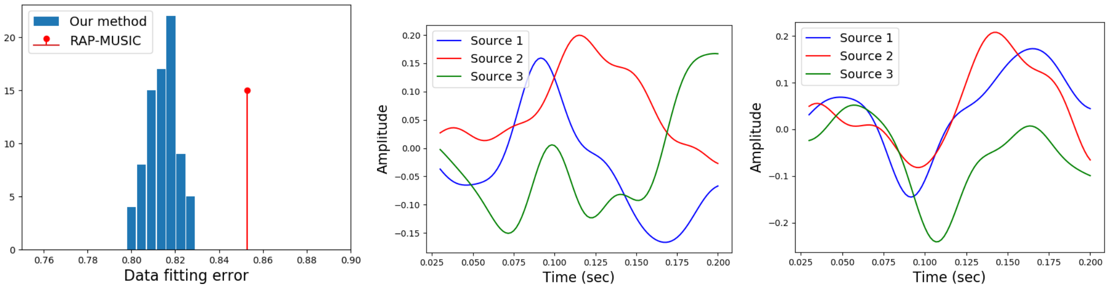
Figure 6.19: Source 2. (a) Best projection error by candidate. First 4 candidates have significantly lower error and are chosen for further analysis. (b) Error as a function of a region size for selected candidates. (c) Upper bound of candidates location on the cortex. Lighter is the shade of red, lower is the error of corresponding region. (d) Topographies associated to the best nodes of candidate growing regions.



(a) (b) (c) (d)

Figure 6.20: Source 3. (a) Best projection error by candidate. First 4 candidates have significantly lower error and are chosen for further analysis. (b) Error as a function of a region size for selected candidates. (c) Upper bound of candidates location on the cortex. Lighter is the shade of red, lower is the error of corresponding region. (d) Topographies associated to the best nodes of candidate growing regions.

Figure 6.21: Cortical representation of the RAP-MUSIC solution (yellow spheres) on top of the regions, found by our method (red, green and blue regions). Different colors correspond to different sources.



(a) **(b)** **(c)**

Figure 6.22: (a) Distribution of triplets' data fitting error (blue histogram) and the error obtained with RAP-MUSIC (red indicator). (b) Source signals found with our method. (c) Source signal found with RAP-MUSIC.

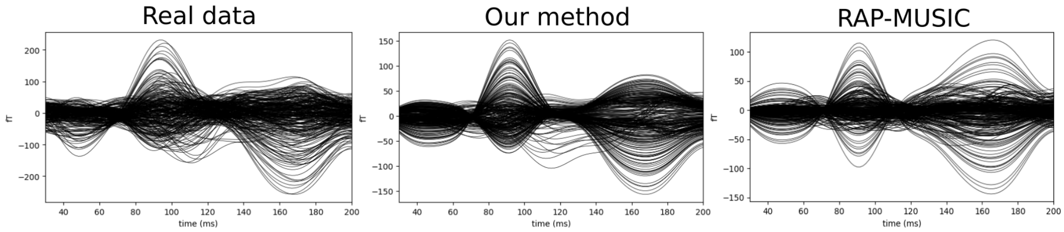


Figure 6.23: The real signal at sensor level and the signal reconstructed with our and RAP-MUSIC methods.

Bibliography

Brainstorm software. Brainstorm software. URL https://neuroimage.usc.edu/brainstorm/.

Steven Brown, Michael J. Martinez, and Lawrence M. Parsons. Music and language side by side in the brain: a pet study of the generation of melodies and sentences. *The European journal of neuroscience*, 2006.

Rahul S. Desikan, Florent Ségonne, Bruce Fischl, Brian T. Quinn, Bradford C. Dickerson, Deborah Blacker, Randy L. Buckner, Anders M. Dale, R. Paul Maguire, Bradley T. Hyman, Marilyn S. Albert, and Ronald J. Killiany. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, 2006. ISSN 10538119. doi: 10.1016/j.neuroimage.2006.01.021.

Alexandre Gramfort, Martin Luessi, Eric Larson, Denis A. Engemann, Daniel Strohmeier, Christian Brodbeck, Roman Goj, Mainak Jas, Teon Brooks, Lauri Parkkonen, and Matti Hämäläinen. MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, 2013. ISSN 1662453X. doi: 10.3389/fnins.2013.00267.

MNE tutorial page. Brainstorm auditory tutorial dataset. URL https://mne.tools/stable/auto_tutorials/sample-datasets/plot_brainstorm_auditory.html#tut-brainstorm-auditory.

Ernst Niedermeyer and F.H.Lopes Da Silva. *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, Fifth Edition*. Lippincott Williams and Wilkins, fifth edition, 2004. ISBN 0-7817-5126-8.

A. C. Papanicolaou, R. L. Rogers, S. Baumann, C. Saydjari, and H. M. Eisenberg. Source localization of two evoked magnetic field components using two alternative procedures. *Experimental Brain Research*, 80(1):44–48, Apr 1990. ISSN 1432-1106. doi: 10.1007/BF00228845. URL https://doi.org/10.1007/BF00228845.

Gerard B. Remijn, Emi Hasuo, Haruna Fujihira, and Satoshi Morimoto. An introduction to the measurement of auditory event-related potentials (erps). *Acoustical Science and Technology*, 35(5):229–242, 2014. doi: 10.1250/ast.35.229.

Francois Tadel, Sylvain Baillet, John Mosher, Dimitrios Pantazis, and Richard M Leahy. Brainstorm: A user-friendly application for meg/eeg analysis. *Computational intelligence and neuroscience*, 2011:879716, 04 2011. doi: 10.1155/2011/879716.

M. A. Uusitalo and R. J. Ilmoniemi. Signal-space projection method for separating meg or eeg into components. *Medical and Biological Engineering and Computing*, 35(2):135–140, Mar 1997. ISSN 1741-0444. doi: 10.1007/BF02534144. URL `https://doi.org/10.1007/BF02534144`.

David Woods. The component structure of the n1 wave of the human auditory evoked potential. *Electroencephalography and clinical neurophysiology. Supplement*, 44:102–9, 02 1995.

General conclusion

Contributions

In this work we have made two main contributions.

A fast algorithm for approximation of the EEG forward problem solution.

We provided a method which is able to approximate a solution of the EEG forward problem, i.e. approximate the lead field matrix, for different head tissue conductivity configurations. It requires to compute only a small number of exact solutions for so-called support points. Our algorithm also provides an optimal way to select these support points. We analysed some theoretical properties of our method, such as its computational complexity, error convergence rate, and also tested its empirical performance in few ways.

We showed that the approximation error decreases very quickly with the number of support points, and that a small number of support points is enough to achieve a good approximation.

We also demonstrated the usefulness of the method in a realistic context of conductivity estimation from EEG data, both simulated and real, using a simple dipole fitting method. As expected, a relatively small number of precomputed matrices provide results which are similar but remarkably faster when compared to using exact matrices.

A clustering-based extension of state-of-art MUSIC algorithm to solve M/EEG inverse problem.

We provided an algorithm for solving M/EEG inverse problem based on the assumption that the M/EEG signal is generated by a set of spatially distinct and extended cortical regions. A region is assumed to have a spatially uniform amplitude. The regions also supposed to be temporally unsynchronised, even though they might be correlated.

Unlike the majority of other inverse methods, which seek a unique solution of the problem, our algorithm is able to provide several spatially distinct candidates for a solutions, which fit the data with similar accuracy. It also provides estimating of the spatial extent of these candidates. In fact, our approach is based on the idea that it is more desirable to obtain a few false positives solutions together with the true positive, than to obtain a single false positive solution.

We tested our method on simulated and real MEG data and showed that several spatially distinct source configurations can, indeed, explain the data with similar accuracy. Our method is able to provide several candidates for the solution.

Some works are not listed in details in this manuscript. In the work with Samuel Deslauriers-Gauthier we considered the MEG source model constrained by white matter streamlines, obtained using diffusion magnetic resonance. In the work of Isa Costantini et al. I provided an expertise about sparse regularization using Least-Angle Regression for LASSO (LARS) method. It was used to deconvolve the blood oxygenated level dependent (BOLD) signal to recover underlying neurons activations and their dynamics. The mentioned works can be found in Appendix "Contributions outside the scope of this thesis".

In addition to research work, I was one of the initiators and developers of the simulation module for the MNE-Python software.

Perspectives

Both contributions of this work have the global objective of improving the accuracy of the M/EEG source localization. With the help of our lead field approximation algorithm, it becomes computationally efficient to estimate head tissue conductivities from the EEG data, which is an important task, because the solution of the EEG forward problem is sensitive to some tissue conductivity values. Using more accurate conductivity values, in return, improves the performance of the source reconstruction.

Further theoretical analysis of the proposed method can be done in future work. In particular, the theoretical error convergence rate and the error bounds outside the sampling mesh should be studied more. As for the application perspectives, on should consider the integration of conductivity estimation with the stat-of-art inverse methods – (T)RAP-MUSIC, Beamforming, MNE, etc.

Our CLUB-MUSIC method aims not only to find the active cortical sources which explain the M/EEG data, but also to provide several candidates for each of them. In this respect, our approach differs widely from the majority of the state-of-art methods of M/EEG source localization. In our work, we show that uniqueness of the solution is not always a desirable property of a M/EEG source localizer.

For example, considering several candidates for the source location can be crucial in epilepsy, where source reconstruction is used in pre-surgical planning. For a clinician, it might be possible to determine which of the candidates, found by our method, is most likely to be the true epileptic source. Apart from M/EEG, there usually are other data which guide a clinician, such as structural MRI, physiological symptoms, etc. A single solution provided bey a M/EEG inverse method might lead to a false localization, which is unacceptable in this context of epilepsy.

Several aspects of our method, however, need further research. The most important of them are the choice of the regularization parameter and thresholds for source selection. At the same time, there is a room for further generalizations of our method. Thus, one might think about constraining regions to grow inside the parcels of a cortical atlas, or to include structural connectivity (estimated from dMRI) to the model.

Lastly, the natural idea for a future work is to combine both methods presented in this thesis into the same framework. More precisely, it would be reasonable to estimate the tissue conductivities and the active cortical regions at the same time using the CLUB-MUSIC and a fast lead field approximation.

Contributions outside the scope of this thesis

White matter fiber bundles as a source model in the MEG inverse problem

Kostiantyn Maksymenko, Théodore Papadopoulo, and Samuel Deslauriers-Gauthier

Université Côte d'Azur, Inria, France

Introduction: The magnetoencephalography (MEG) inverse problem, i.e. recovering brain activity from MEG measurements, is ill posed and additional hypotheses are needed to constrain the solution space. A common approach is to constrain the problem by taking into account the subject's anatomy. In this work, we introduce an approach which considers white matter streamlines, obtained using diffusion magnetic resonance (MR) imaging, as a source model for the MEG forward problem. To simplify the model and reduce the computational complexity we regrouped similarly shaped streamlines into bundles. The MEG data associated with a single bundle activity was simulated. The objective was to fit simulated data for each bundle and to analyze the data fitting error.

Methods: Based on diffusion and anatomical MR images, we computed streamlines using anatomically constrained tractography [1]. Because of the anatomical constraint, the streamline end-points were located on the interface between white and gray matter and therefore provided anatomical support for the MEG source space. In order to reduce the number of sources to deal with in the MEG inverse problem, we applied a clustering algorithm to regroup streamlines into bundles [2]. This resulted in 1266 bundles (instead of 1.2 million streamlines) which represented connections between different cortical regions. We then computed MEG forward operator (so-called lead field matrix) using OpenMEEG. We used the streamlines' end-points as sources. The head and electrodes models were computed based on the MRI and real MEG sensor positions of the subject.

Two important assumptions were made to simplify the model. First, all sources associated with a bundle have the same activity. This allowed us to sum the lead fields (the columns of the lead field matrix) of the sources associated with one end of a bundle. Second, we assumed a delay between activities of the first and the second end of a bundle. This delay was estimated as the average length of a bundle divided by 6 m/s [3]. To simulate the measurements, we used an autoregressive model:

$$y(t) = l_1 \cdot x(t) + l_2 \cdot x(t - \delta) \tag{1}$$

where $y(t)$ is the measured MEG signal, l_1 and l_2 are lead fields of the two bundle ends, δ is the estimated delay, and $x(t)$ is the bundle activity. For a selected bundle, we simulated the source activity pattern $x(t)$ then the measurements $y(t)$ using (1). We then fitted the data $y(t)$ for each bundle independently, solving a least square inverse problem.

Results: We applied our technique to the subject 100307 from the Human Connectome Project (HCP) dataset. Figure 1A illustrates the simulated MEG signal. Solving the inverse problem for each bundle independently we obtained a data fitting error. We then transformed the error as $\hat{e}_i(t) = \max_j(e_j(t)) - e_i(t)$, $\forall i \in N$ where N is the number of bundles. This value represents the accuracy of each bundle. Figure 1B illustrates the accuracy of the bundles at different moments in time. The more intense the red color, the better the data fit. As expected, the simulated bundle had the best accuracy.

Discussion and Conclusion: We proposed a new approach to constrain MEG inverse problem based on the white matter fiber bundles of a subject. We computed the accuracy of each bundle independently of others which seems to be a good criterion in the case of single bundle activity. This approach has a natural limitation if several bundles are active simultaneously. But even in this case our approach can provide some information about the process and can be used as a preprocessing before applying more complex inverse problem methods [4].

Has natural limitations, but can be used as a preproccesing step for more complex models

- input "info" instead of "raw" - noise and artifacts are generated separately - fwd object as a default input, but keep a possibility to compute forward problem inside

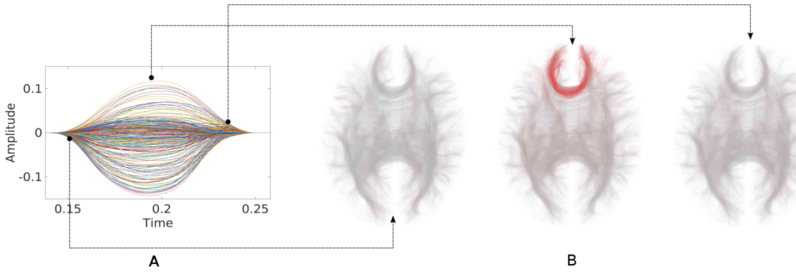


Figure 1: Simulated MEG signal based on one bundle activity. A) Time series represent a signal of each channel. B) We also show accuracy of each bundle at the beginning, the peak, and the end of the signal.

References

[1] Girard, G., Whittingstall, K., Deriche, R., Descoteaux, M.: *Towards quantitative connectivity analysis: reducing tractography biases*, NeuroImage, 98, 266-278, 2014.

[2] Côté M-A., Garyfallidis, E., Larochelle, H., Descoteaux, M.: *Cleaning up the mess: tractography outlier removal using hierarchical QuickBundles clustering,* Proceedings of: International Society of Magnetic Resonance in Medicine (ISMRM), Toronto, Canada, 2015.

[3] Tomasi, S., Caminiti, R., Innocenti, G.: *Areal differences in diameter and length of corticofugal projections*, Cerebral Cortex, 22, 1463-1472, 2012.

[4] Deslauriers-Gauthier, S., Lina, J-M. , Butler, R., Bernier, P-M., Whittingstall, K., Deriche, R., Descoteaux, M.: *Inference and Visualization of Information Flow in the Visual Pathway using dMRI and EEG,* MICCAI 2017 Medical Image Computing and Computer Assisted Intervention, Sep 2017, Québec, Canada.

fMRI Deconvolution via Temporal Regularization using a LASSO model and the LARS algorithm

Isa Costantini, Patryk Filipiak, Kostiantyn Maksymenko, Samuel Deslauriers-Gauthier, Rachid Deriche
Inria Sophia Antipolis-Méditerranée, Université Côte d'Azur, France

***Abstract*— In this work we propose a novel approach to deconvolve the blood oxygenated level dependent (BOLD) signal to recover underlying neurons activations and their dynamics.**

I. INTRODUCTION

Deconvolution methods are used to denoise the blood oxygen level-dependent (BOLD) response, the signal that forms the basis of functional MRI (fMRI) [1]. In this work we propose a temporal regularized deconvolution of BOLD fMRI signal with the least absolute shrinkage and selection operator (LASSO) model, solved using the Least-Angle Regression (LARS) algorithm.

II. METHODS

We implemented the deconvolution exploiting the sparsity of the innovation signal s(t) [2], and we built a LASSO model:

$$s^* = \underset{s}{\operatorname{argmin}} \left\{ \frac{1}{2n} \left\| y - H I_\alpha s \right\|_2^2 + \lambda \| s \|_1 \right\}$$

where n is the signal length, H represents the hemodynamic response function [3] and λ is the regularization parameter. I_α *is the exponential accumulation function given by:*

$$I_\alpha(z) = S \left[\frac{e^{-\alpha} z^{-1}}{(1 - e^{-\alpha} z^{-1})^2} - \frac{e^{-\alpha} z}{(1 - e^{-\alpha} z)^2} \right] \frac{1}{(1 - z^{-1})^2}$$

such that the activity-inducing signal $u(t) = I_\alpha * s(t)$, *normalized by the factor S. We chose α = 0.75 experimentally within the range [0, 3]. We minimized the objective function using LARS [4], that outputs all λs of interest and their associated solutions. We used the L-curve to estimate the* λ^* *corresponding to the optimal solution* s^*. *Then, as in [2], we scaled a 3D activation map computed with FSL in the range [0, 3], with a 2-mm isotropic resolution (Fig.1b). We multiplied it by two block-type signals of 200s, u(t): A) with 4 onsets, and B) with one long onset. We corrupted the signals with model and block-type noise, we convolved them with H and we added Gaussian noise thus simulating the fMRI time series y(t). We solved the inverse problem and we recovered* $u^*(t) = I_\alpha * s^*(t)$ *as described above. Finally, to evaluate the results, we computed the root of the mean square errors (MSE) and standard deviation (STD) between u(t) and* $u^*(t)$ *averaged among the voxels belonging to the grey matter masked activation. We compared our results with those obtained with the temporal regularization implemented in the Total Activation (TA) toolbox [1]. We tested the above procedure on the preprocessed task-fMRI image of one subject taken from the Human Connectome Project (HCP) database [5]. The reconstructed* $u^*(t)$ *were averaged in a ROI of 6 × 6 × 6* mm^3 *centered in the Brodmann Area 4p (MNI coordinates: 62, -14, 30).*

III. RESULTS

Table I shows that the MSEs±STDs change for different peak-SNRs (pSNRs) and that they are lower than the ones obtained using the TA toolbox. Fig.1.a and c show examples of the reconstructed activity inducing signal using our approach (u^*) *and the approach in [1]* (u^*_{TA}).

TABLE I. SUMMARY OF ROOTED *MSEs* AND *STDs*

Activation		**A**			**B**		
pSNR [dB]		5.17	4	3.98	7.64	5.94	5.12
OUR *α = 0.75*	*RMSE*	**0.11**	0.22	0.18	**0.05**	0.19	**0.14**
	STD	**0.13**	0.16	0.14	**0.05**	0.11	**0.08**
TA	*RMSE*	0.24	0.29	0.3	0.19	0.24	0.25
	STD	0.31	0.32	0.33	0.23	0.24	0.26

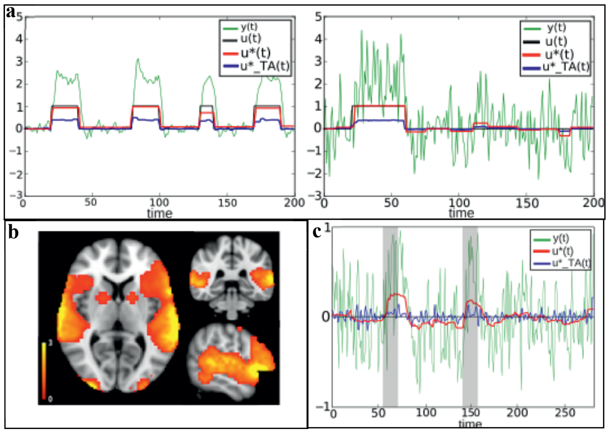


Fig. 1: (a) Reconstructed signal u* obtained with our approach (red) and the TA (blue) superimposed on the activation (black) and fMRI signal (green). The plots are related to different activations and pSNRs: activation A, pSNR = 5.17 dB (left); activation B, pSNR = 5.12 dB (right). (b) Activation map. (c) Results of HCP data. Same legend of panel (a); the grey areas represent the duration of the tongue movements.

IV. DISCUSSION & CONCLUSION

Our findings show that the joint use of LARS and the L-curve for solving our optimization problem allowed us to choose the optimal λ^* *and related solution among all those outputted by the algorithm. Thus, we decreased the computation time and avoided a need of defining λs a priori, allowing to improve brain dynamics recovery in future clinical application.*

REFERENCES

[1] F. I. Karahanolu, *Total activation: fMRI deconvolution through spatio-temporal regularization.* Neuroimage, 2013.
[2] Y. Farouj, *Regularized spatiotemporal deconvolution of fMRI data using gray-matter constrained total variation,* ISBI 2017.
[3] I. Khalidov, *Activelets: Wavelets for sparse representation of hemodynamic responses,* Signal Processing, 2011.
[4] B. Efron, *Least angle regression,* The Annals of statistics, 2004.
[5] D.C. Van Essen, *The WU-Minn HCP,* NeuroImage, 2013.

ACKNOWLEDGMENTS - This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (ERC Advanced Grant agreement No 694665 : CoBCoM - Computational Brain Connectivity Mapping).